Design and Development of the Wage and Investment Compliance Burden Model

Michael Stavrianos and Arnold Greenland

Abstract

Each year, individuals in the United States submit more than 120 million tax returns to the Internal Revenue Service (IRS). The IRS uses the information in these returns, recorded on hundreds of distinct forms, to administer a tax system whose rules span thousands of pages. Managing such a complex and broad-based tax system is costly; in fiscal year 2000, the budget of the IRS exceeded $8 billion. But these costs represent only a fraction of the total burden of the tax system. Equally, if not more burdensome, is the time and money that citizens spend in order to comply with tax laws and regulations. The IRS estimates that taxpayers spend nearly 6 billion hours each year on tax compliance activities, such as record keeping, tax planning, form completion, and form submission.

The IRS is working with PwC Consulting to develop an improved methodology for measuring and modelling the burdens imposed on the public by the Federal tax system. This study, and the resulting model, will assist the IRS in its mission to provide taxpayers with top quality service—and it will help policymakers understand the full impact of changes in tax law. The purpose of this conference report is to summarize the work conducted during the first phase of the taxpayer burden project, which focused on Wage and Investment Taxpayers.

Author Bios

Mr. Michael Stavrianos, a Principal Consultant at PwC Consulting, has more than eight years of experience in data analysis, statistical modeling, survey research, and project management. He has particular expertise in policy applications of microsimulation modelling techniques and in the use of microsimulation models to provide decision support for government clients. Mr. Stavrianos led the development and implementation of the W&I Compliance Burden Model. He holds a Masters of Public Policy from Georgetown University and a B.A. in Statistics from Washington University.

Dr. Arnold Greenland, a Managing Director at PwC Consulting, has more than 25 years of experience managing analytically oriented research projects for both public and private sector clients. He has served as project director on all of PwCC's burden modelling engagements, and has managed projects in the areas of operations research, business process modeling and simulation, business forecasting, network optimization, and systems development. He also served as a lecturer at the George Washington and George Mason Universities on operations research. Dr. Greenland holds a Ph.D. in Mathematics from the University of Rochester and a B.A. in Mathematics from Case Western Reserve.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

I. INTRODUCTION

Each year, individuals and businesses in the United States submit more than 200 million tax returns to the Internal Revenue Service (IRS). The IRS uses the information in these returns, recorded on hundreds of distinct forms, to administer a tax system whose rules span thousands of pages. Managing such a complex and broad-based tax system is costly—in fiscal year 2000, the budget of the Internal Revenue Service exceeded $8 billion. But these costs represent only a fraction of the total burden of the tax system. Equally, if not more burdensome, is the time and money that citizens spend in order to comply with tax laws and regulations.

The IRS estimates that taxpayers spend approximately 6 billion hours each year on tax compliance activities, such as tax planning, record keeping, and form completion.[1] This estimate is based on a model developed in 1984 by the IRS and Arthur D. Little, Inc. (ADL). Over the years, the ADL model has become outdated on several counts.[2] In particular, the following criticisms have been raised:

- **The model is based on survey data collected more than 15 years ago**. Since that time, changes in tax regulation, tax administration, and tax preparation methods have changed the amount and distribution of taxpayer burden.

- **The model measures only the *paperwork* burden of forms and regulations.** Paperwork burden, as defined by the Paperwork Reduction Act of 1980, is only one component of total taxpayer burden.

- **The model has limited predictive power**. Burden is estimated using only a few simple determinants, such as the number of line items on a form and the length of instructions. As a result, the model can simulate only a narrow range of policy changes. Moreover, the model does not accurately reflect variation in burden across taxpayers.

- **Burden is measured by hours alone, not dollars**. Dollar estimates are needed to support policy decisions, particularly regarding the tradeoff between IRS budget, tax revenue, and taxpayer burden.

The IRS is working with PwC Consulting and the Taxpayer Burden Working Group[3] to develop an improved method for measuring taxpayer compliance burden. This new method will be a valuable tool in the IRS's ongoing effort to improve customer service, as well as for policy makers to understand the full

---

[1] Fiscal Year 2001 Information Collection Budget of the United States Government, Office of Management and Budget, Office of Information and Regulatory Affairs.

[2] For a more detailed critique of the ADL model, see Section II.A of Phase A Report - Final (July 22, 1999).

[3] The Taxpayer Burden Working Group was convened to guide the design and implementation of IRS's burden research. The core Working Group is composed of stakeholders from IRS, the Office of Tax Analysis (Department of Treasury), and the Office of Management and Budget.

effect of tax rule changes. In particular, it will help the IRS understand the burdens placed on its customers by the federal tax system—its laws, its administration, and changes to those factors.

I.1 PROJECT OBJECTIVES

The ultimate goal of this multi-phase project is to provide the IRS with an improved methodology for measuring and modeling the burdens imposed on the public by the Federal tax system. This study, and the resulting model, will assist the IRS in its mission to provide taxpayers with top quality service. Specific objectives related to this overarching goal include:

- **Define Taxpayer Burden.** Build consensus on a definition that addresses the various types of burden surrounding the tax system, but recognizes that some of these burdens are not attributable to tax compliance.
- **Measure the Level of Taxpayer Burden.** Develop a measurement approach that provides detailed and accurate measures of taxpayer burden, in terms of both time and out-of-pocket costs.
- **Forecast Changes in Burden.** Develop a model that allows policymakers and policy analysts to estimate the burden impact of changes in tax policy, tax system administration, or other factors.
- **Build a Flexible Framework.** Design a measurement and modeling methodology that is flexible enough to accommodate diverse segments of the taxpaying population—and that can produce burden estimates for a variety of purposes.

This report documents the work conducted during the first phase of the burden project, aimed at developing a burden model for Wage and Investment (W&I) taxpayers.[4] The W&I service market includes more than 90 million individual taxpayers and thus accounts for a large share of total burden. The project's second phase, currently underway, will expand the model to include Self-Employed (SE) taxpayers—thus yielding burden estimates for all individual taxpayers. Subsequent phases of the project will address the measurement and modeling of burden for other taxpayer segments, incorporating lessons learned from the individual-taxpayer prototype model.

I.2 DEFINING TAXPAYER BURDEN

Taxpayer burden includes several components. The greatest of these burdens is generally tax liability; any tax system imposes a burden on taxpayers equal to the amount of the tax. Additionally, any tax other than a lump sum levy imposes "excess burdens," which include:

[4] Wage and Investment taxpayers have been defined by the IRS as individuals who file Form 1040, 1040A, or 1040EZ, but do not file Schedule C, Schedule E, Schedule F, or Form 2106. For convenience, this report uses the term "taxpayer" to refer to the tax-filing unit.

(1) Out-of-pocket expenses incurred to comply with the tax system

(2) Time foregone to comply with the tax system

(3) Psychological Costs

(4) Net Efficiency Costs resulting from distortions in income or consumption patterns due to the tax system (e.g., purchase vs. rent housing)

Taxpayers can affect the allocation of burden among tax liability and the four excess burden categories through their behavior and reactions to the tax system. For example, taxpayers can spend more time and money (components of excess burden) on tax planning in order to reduce the amount of tax they owe (the tax liability component of burden). They can also spend more money using a paid preparer (one component of excess burden) in order to spend less of their own time on taxes (another component of excess burden). Any total measure of taxpayer burden must include each of the components, and recognize the interaction and tradeoffs between those components.

Different types of analytical models are used currently to estimate the major components of total taxpayer burden. The tax liability component of total taxpayer burden is estimated by the U.S. Treasury Department using microsimulation models based on tax return information. The efficiency cost component of excess burden is generally measured using macro-econometric models that are structured at the regional or national level. Psychological costs, which are not captured in any of the other models, are generally considered to be beyond the practical ability of computer models to estimate.[5]

While all the components of total burden are important, the focus of this study is **taxpayer compliance burden**—the time and money that taxpayers spend to comply with the federal income tax system.[6] The advantages of this definition include: (1) it is an intuitive concept of compliance burden, (2) it eliminates redundancies and potential inconsistencies across burden components (e.g. avoids double counting of the part of taxpayer burden picked up in revenue estimates), and (3) it is consistent with OMB burden measurement guidelines for the Paperwork Reduction Act.

[5] Although psychological costs are not measured directly, many are captured in one of the other components of the burden measure. For example, if a taxpayer fails to minimize their taxes because they fear the consequences of tax avoidance, psychological costs will be observed as excess taxes. Alternatively, if a taxpayer uses a paid preparer to reduce their stress over tax compliance, psychological costs will be observed as out of pocket costs. The psychological costs that remain (e.g., taxpayer anxiety) are real, but produce no observable behavior. These costs will not be captured in any burden category.

[6] Many activities and costs commonly associated with tax compliance are necessary not only to comply with the federal income tax system, but also for other purposes such as state taxes or loan applications. In cases where a single activity is motivated both by federal tax requirements and by other requirements or interests, the joint costs of the activity must be allocated. A reasonable approach is to designate one set of activities as foundational, and assign all common costs to that set of activities. The definition used in this study treats federal tax requirements as foundational to state tax requirements; and other requirements (e.g., financial planning and reporting) as foundational to all tax requirements.

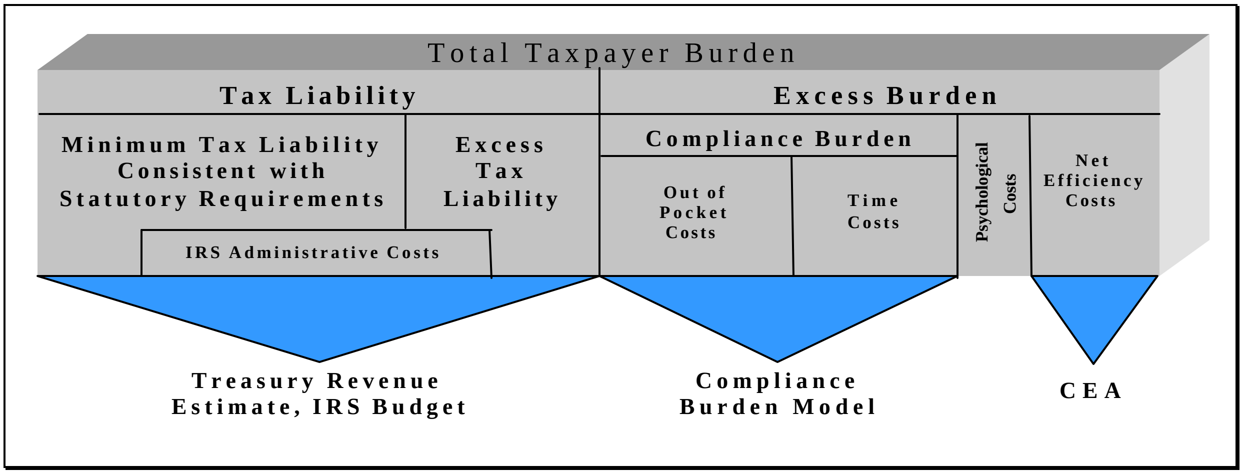
Traditionally, taxpayer compliance burden has been measured in terms of the amount of time (hours) expended, but this definition is incomplete. An estimate of the dollar cost of compliance burden—including a monetized value of taxpayer time—is essential to support decisions regarding the tradeoff between IRS budget, tax revenue, and taxpayer burden. Currently, there is no consensus in the research community regarding the best method for monetizing time. In light of this disagreement, the IRS commissioned a research paper as part of this study to review alternative monetization methodologies as they relate to tax compliance burden.[7]

Finally, it is important to note that the burden *experienced* by wage and investment taxpayers is being measured. In some cases, burdens *produced* by these taxpayers may extend to other agents. For example, an individual transfers compliance costs to his or her employer through completion of a Form W-4 and through the costs of payroll deduction. The network of burden widens further when one considers banks, brokerages, mortgage companies, and other groups that are required to report financial information to the taxpayer and to the IRS. These burdens, induced by individuals but experienced by others, will be measured at the locus of the burden.

No single measure of burden is appropriate for all purposes. IRS needs the flexibility to combine different components of burden to construct measures that are suitable for a variety of purposes. The segmented definition of total burden presented here allows for this type of aggregation. Together, the models used by the IRS, Treasury, and others will capture all of the measurable components of total taxpayer burden. ***Figure I-1*** illustrates the various components of total burden and the groups that are working to measure each component.

[7] See Revealed And Stated Preference Estimation of the Value of Time Spent for Tax Compliance (May, 2000), by Dr. Trudy Ann Cameron.

FIGURE I-1: COMPONENTS OF TOTAL TAXPAYER BURDEN



I.3 PURPOSE AND STRUCTURE OF THIS REPORT

The purpose of this report is twofold. The first purpose is to summarize the work conducted during the first phase of the taxpayer burden project, focusing on research methodology and results. The second is to document the key design and implementation decisions made by the Taxpayer Burden Working Group, explain why these decisions were made, and highlight some of the alternatives considered.

In Section II of this report, we describe our approach to modeling taxpayer burden and review the factors that motivated this approach. In Section III, we describe our data collection methodology and provide descriptive statistics for the data sources used to estimate the W&I burden model. In Section IV, we describe our estimation methodology from both a theoretical and empirical standpoint, and evaluate the results of our estimation. In Section V, we describe the development of the production model and explain how the model can be used to forecast the burden impact of a variety of policy or administrative changes.

II. MODELING APPROACH

In this section, we present an analytical framework for modeling taxpayer compliance burden. We begin with a review of high-level model requirements and an explanation of how these requirements motivated the conceptual design of the model. Next, we describe the operational approach that emerged from our conceptual design. Finally, we discuss the three major tasks associated with creation of the burden model—data collection, model estimation, and development of the production model.

II.1 CONCEPTUAL MODEL DESIGN

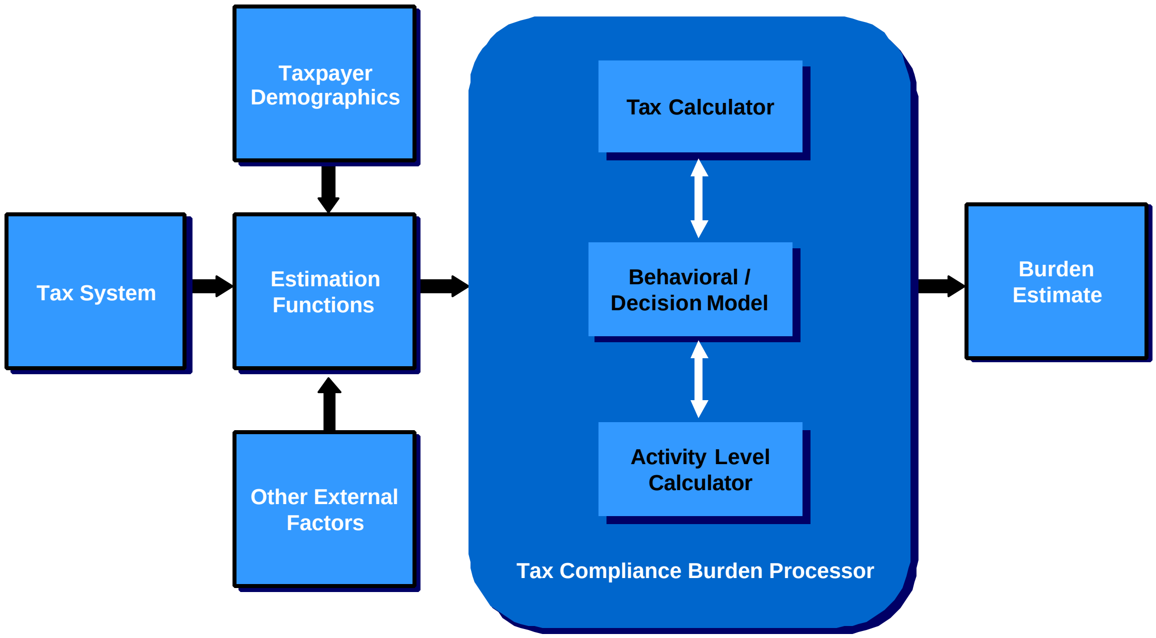
The conceptual framework for a model of taxpayer compliance burden must be robust yet flexible, as the IRS intends to use the model for many purposes. The broad functionality required of the model is evidenced by the following list of requirements—compiled from the Statement of Work and from conversations with stakeholders inside and outside the IRS.

- **Estimate the impact of a wide range of inputs on taxpayer compliance burden** (e.g., changes in tax laws, tax administration, taxpayer demographics, technology)
- **Measure the amount of burden by taxpayer throughout the tax settlement process** and predict changes in taxpayer behavior that affect the level of compliance burden.
- **Disaggregate total burden across several dimensions** (e.g., by type of activity, by type of taxpayer, by type of tax).
- **Utilize parameters and other simulation levers that can be measured or estimated by IRS**. The model should provide the flexibility to alter policy parameters and assumptions about taxpayer behavior in a consistent and straightforward way.

Our selection of a conceptual modeling approach was guided by the requirements specified above, particularly the first three: allowing for a wide range of inputs, estimating burden by taxpayer, and disaggregating total burden across several dimensions. These three requirements led us to conclude that a microsimulation model was the most appropriate tool for generating the desired burden estimates. ***Figure II-1*** illustrates the high level structure of such a model.

Microsimulation is a technique widely used to investigate the impact of public policies by examining the behavior of agents at the micro-level. Microsimulation models operate by taking a representative sample of agents and applying mathematical algorithms to simulate various behaviors and outcomes. By changing the parameters of these algorithms, policymakers can simulate a change in inputs and observe the resulting change in outcomes. Because the model runs on micro-level data, impacts can be measured both in the aggregate and for various subgroups of agents.

FIGURE II-1: CONCEPTUAL MODEL STRUCTURE



Microsimulation is a particularly useful modeling approach in circumstances where, (1) policymakers wish to see how their decisions affect selected subgroups of the population across a number of dimensions, or (2) the policies being simulated involve complex and non-linear interactions with other provisions.

II.1.1 ALTERNATIVE CONCEPTUAL APPROACHES CONSIDERED

A number of conceptual approaches other than microsimulation were considered for the taxpayer burden model, but were rejected by the working group because they did not adequately meet one or more of the key modeling requirements. Among the options considered were the following:

- **Macro-Level Models,** which generate estimates of total burden based on macro-level inputs (e.g., number of taxpayers, GNP, total tax liability) or time-series data (e.g., trends in taxpayer burden, trends in the use of tax software or electronic filing). This approach was rejected because it would not provide estimates of burden for taxpayer subgroups.

- **Cell-Based Simulation Models,** which generally involve an integrated set of macro-level models, each customized for a "cell" of individuals that share some characteristics (e.g., similar levels of tax complexity or income). While cell-based models capture some differences across taxpayer subgroups, they cannot fully account for the complex interactions of tax policy, which involve a wide variety of demographic and income characteristics.

- **Agent-Based Models,** which can be thought of as a subclass of microsimulation models in which the agents interact with one another. These systems are often extremely complex, particularly when the behavior of individual agents is heavily dependent on other agents in the system, or when individual agents are allowed to adapt to their environment. As these characteristics were not considered critical in a model of tax compliance burden, this modeling approach was discarded as unnecessarily complex.

- **Expert Systems,** which apply qualitative information (derived from experts) to individual cases using a pre-determined set of rules. An example of an expert system would be a model that assigned a dollar amount of burden to each tax return based on information obtained from tax professionals. This approach was rejected because it could not easily generate separate estimates of time and money burden, and because the professional fee associated with a tax return is not necessarily a complete or accurate measure of the burden experienced by the taxpayer who prepared the return.

II.1.2 VALIDATION OF THE CONCEPTUAL APPROACH

To validate our conceptual approach and further inform development of the burden model, PwC Consulting conducted focus groups and in-depth interviews with both wage and investment taxpayers and tax preparers.[8] Between February 16, 1999 and February 25, 1999, nine focus groups of individual wage and investment income earners and one focus group of tax preparers were conducted in four different metropolitan areas: St. Louis, Missouri; Seattle, Washington; Houston, Texas; and Philadelphia, Pennsylvania. Through this qualitative research, we validate a number of basic assumptions surrounding our conceptual model of compliance burden, as described below:

- **Taxpayers make tradeoffs between compliance, burden, and tax liability.** Some taxpayers avoid taking legitimate deductions because they do not want to incur the additional burden. These taxpayers may be averse to recordkeeping, or may be concerned that claiming an uncommon deduction—such as the home office deduction—could increase scrutiny from the IRS.

- **Taxpayers generally exhibit inertial behavior.** Taxpayers generally follow the same compliance process each year. There are, however, triggers that often cause taxpayers to change their compliance behavior, including: (1) life changes, such as a divorce or birth of a child; (2) financial changes, such as the purchase of a home or a change in employment status; and (3) tax law changes, such as a change in the treatment of capital gains income.

- **Many factors influence taxpayer burden.** The amount of time and money that taxpayers spend on tax compliance is influenced by many factors, including tax system complexity, taxpayer characteristics, and compliance methods.

- **Preparation method has a major impact on taxpayer burden.** Two fundamental decisions a taxpayer makes with respect to compliance burden are deciding whether to use a paid

[8] For a more detailed discussion of this qualitative research, see Section II.C of Phase A Report - Final (July 22, 1999).

professional and deciding whether to use tax software. The choice of preparation method influences pre-filing, filing, and post-filing activities.

- **Submission method also influences burden, but to a lesser degree than preparation method.** This is especially true among taxpayers who Tele-File. Tele-Filing is a primary driver of burden throughout the tax preparation process, with the exception of a few pre-filing activities.

- **The more experience one has in doing taxes the less of a burden it becomes.** As taxpayers become more experienced, they spend less time learning about and completing their taxes.

II.2 OPERATIONAL MODEL DESIGN

The process of transforming the conceptual design into an operational design involved a number of detailed decisions. These decisions centered around three main topics—model functionality, data inputs, and simulation levers.

II.2.1 MODEL FUNCTIONALITY

The conceptual design presented above provides a high-level description of model functionality. It states that the model should provide estimates of time and money burden at a taxpayer level and accommodate a variety of "what-if" scenarios. Taking this as a starting point, PwC Consulting collaborated with the working group to identify specific scenarios that the model should be able to handle. Two key findings emerged from this discussion. First, stakeholders are most interested in scenarios that involve changes in **tax policy** or **tax system administration**. Second, stakeholders recognize that **tax compliance methods** (e.g., preparation method, submission method, use of IRS services) are important determinants of burden, and are interested in the ability to simulate changes in these methods.

These findings led us to conclude that the burden model should include two major functional components in addition to the burden estimation component: (1) a tax engine to predict the filing outcomes that result from the interaction of taxpayer characteristics, tax policy, and tax system administration; and (2) a decision module to predict taxpayer choices regarding compliance methods.

II.2.2 DATA INPUTS

To support the estimation of relationships in the burden model, a data set is needed that contains taxpayer-level information on time and money burden, as well as the factors that generate that burden (e.g., compliance methods and filing outcomes). Recognizing these data needs, our operational design called for the development of an integrated taxpayer-level data file through the following four steps: (1) obtain a nationally representative sample of primary taxpayers who filed W&I returns in Tax Year 1998 (TY98),

(2) administer a survey to a subset of these taxpayers to measure the time and money burden associated with their TY99 tax returns, (3) obtain TY98 and TY99 tax return data for each member of the taxpayer sample, and (4) merge the TY98 and TY99 tax return data to each record in the survey data set.

There is an important distinction between the data source (described above) that is used to estimate the burden model and the data source that is used to run the model—often referred to as a "production file." For model estimation, the data file must provide a comprehensive picture of taxpayer decisions, activities, and burden. This file, while rich in detail, is limited in size due to the cost of collecting survey data, and therefore may be too small to allow microsimulation analysis of burden among small subgroups of taxpayers. The estimation file is further limited in that the survey sample may not accurately reflect the demographic characteristics of the W&I population as time passes. To overcome these limitations, the operational design recommended that a larger sample of individual tax returns—derived from an IRS administrative source—be used as the model's production file.

II.2.3 SIMULATION LEVERS

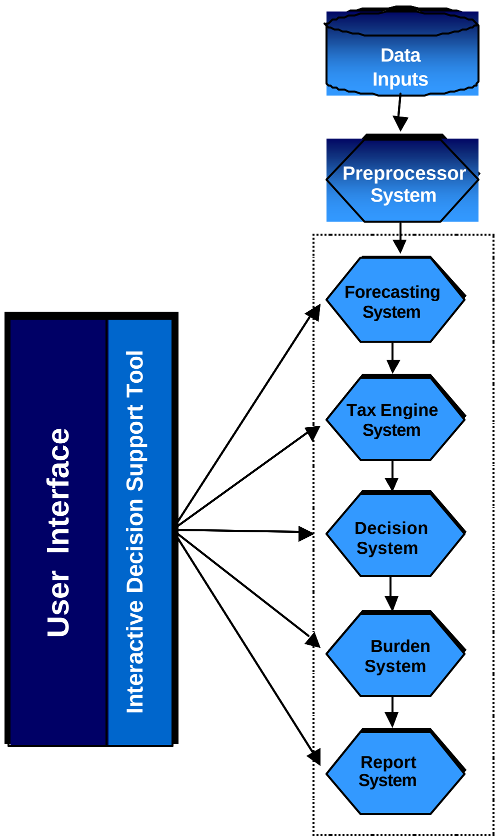
Two of the key functional objectives established at the outset of this project are, to a degree, at odds with one another. The first objective is that the burden model should allow analysts to simulate the change in burden resulting from changes to a wide variety of inputs (e.g., taxpayer demographics, economic trends, tax policy or administration). The diversity of these inputs pushes us towards expanded model functionality and a wider array of simulation levers. The second is that the burden model should have a simple interface that allows users to alter policy parameters or behavioral assumptions in a straightforward way. This objective of simplicity argues for parsimony in the selection of simulation levers.

In operationalizing the model, the working group devoted considerable effort to the selection of appropriate simulation levers. A number of factors were considered in this process, including the importance of the lever in influencing burden, the user's ability to obtain accurate input information, and the model's ability to predict the impact of the lever. Collectively, these decisions resulted in a user interface that, (1) provides a diverse set of simulation levers, (2) can be used easily by non-programmers, and (3) assumes that users are analytically sophisticated and familiar with the design of the model.

II.3 IMPLEMENTING THE OPERATIONAL MODEL DESIGN

Figure II-2 shows the high-level operational design for the W&I taxpayer burden model, along with a brief description of each model component.

FIGURE II-2: OPERATIONAL MODEL STRUCTURE



User Interface / IDST: A graphical user interface that gives users access to simulation levers and helps users create what-if scenarios

Data Inputs: Input to the burden model, describing taxpayer demographics, filing outcomes, and other key characteristics

Preprocessor System: A data preparation module that integrates tax return data with survey data and imputes selected data elements

Forecasting System: A data-aging module that adjusts weights and income/expense amounts based on user specifications

Tax Engine System: An enhanced tax calculator that evaluates tax rules and taxpayer characteristics to determine filing requirements and filing outcomes

Decision System: A simulation model that predicts taxpayer decisions regarding compliance methods (preparation method, submission method, use of taxpayer services)

Burden System: A simulation model that predicts time and out-of-pocket burden based on taxpayer characteristics, filing outcomes, and compliance methods

Report System: A report generator that tabulates the distribution of burden across various dimensions and creates output data files

To implement this operational design, we proposed a three-phase project, comprising:

- **Data Collection,** in which we collect survey data and merge it with IRS administrative data, yielding a comprehensive estimation data file that contains information on taxpayer characteristics, activities, and burden.
- **Estimation,** in which we use econometric and statistical techniques to identify and investigate relationships between taxpayer characteristics, taxpayer decisions, and compliance burden.
- **Model Development,** in which we develop a software tool that, (1) applies the estimated relationships to a production data file, (2) gives users access to simulation levers, and (3) generates summary reports that describe burden under different scenarios.

The remainder of this report is organized around these three phases. Section III deals with data collection, Section IV with estimation, and Section V with model development. Before proceeding, though, it is important to note that these three phases can and should continue throughout the life of the model. Indeed, two types of “model maintenance” were specifically recommended as part of the operational design. First, the production data file should be updated annually with more recent taxpayer data. This strategy is particularly appealing since it dovetails with data file development processes already performed by various groups within IRS. Second, targeted data collections should be conducted to update and enhance the input data. Incorporating new data in this way will allow the model to capture aspects of burden that become more prominent in the future without wholesale replacement of the initial taxpayer survey.

III. DATA COLLECTION

As stated in Section I.1, the ultimate goal of this project is to provide the IRS with an improved methodology for measuring and modeling the burdens imposed on the public by the Federal tax system. To support the model estimation tasks, PwC Consulting collected data from W&I taxpayers to understand the time and out of pocket expenses incurred to comply with federal tax rules and regulations.

During a six-month data collection period (May 1, 2001 to October 31, 2001) respondents completed either a twenty-minute telephone interview or a ten-page questionnaire. Respondents were asked questions about a variety of tax related activities, as well as questions about the time and money they spent to comply with federal tax rules and regulations.

III.1 SAMPLE DESIGN

The specifications of the sample design balanced two main objectives. First and foremost, the sample must contain a sufficient number of cases to support the estimation of empirical models of taxpayer behavior and burden. Secondly, the sample must be distributed efficiently in order for the estimates from the sample to be statistically reliable.

The IRS and PwC Consulting selected the Return Transaction File (RTF) as the sample frame for this phase of the study and decided that the tax return would be the sampling unit.[9] The tax return, rather than the taxpayer, was chosen as the sampling unit because the tax return is the unit of observation in the RTF data file, and because it is more practical to model burden at the tax return level.

The W&I population was segmented by three different variables: (1) form and schedule type, (2) preparation method, and (3) adjusted gross income (AGI). These three variables were chosen for stratification because of their expected correlation with taxpayer burden. In addition, by defining appropriate cut points for each stratification variable, we were able to ensure representation of small, but important, subgroups of taxpayers. The stratification variables and categories are presented in ***Table III-1***.

[9] The Returns Transaction File—a subset of IRS's Master File—contains detailed information from each tax return filed in a given year.

TABLE III-1: SAMPLE VARIABLES FOR STRATIFICATION

Form and Schedule Type	Adjusted Gross Income	Preparation Method
Form 1040 with no Schedule A and no Schedule D filed, not exclusive	Less than $30,000	Paid Professional – Used a paid professional, including if the paid professional used software
Form 1040 with a Schedule A, but no Schedule D filed, not exclusive	Between $30,000 and $70,000	Tax Software – Self-preparers who used tax software
Form 1040 with a Schedule D filed, not exclusive	More than $70,000	Self Preparation – used no software, including Tele-Filers
Form 1040A, not exclusive		
Form 1040EZ and Tele-File		

Although the stratification variables define 45 unique cells, in some instances, cells were combined to reflect areas where it was expected that the population was relatively homogeneous in terms of burden. This resulted in 20 cells. For purposes of estimation, we established a target of 300 completed data records in each cell. As a result, our goal was to complete 6,000 interviews with W&I taxpayers.

The distribution of the 20 cells as well as the population distribution for this universe, based on the 1998 tax year RTF, is presented in ***Table III-2***.

TABLE III-2: POPULATION OF 1998 WAGE & INVESTMENT TAXPAYERS
(in thousands)

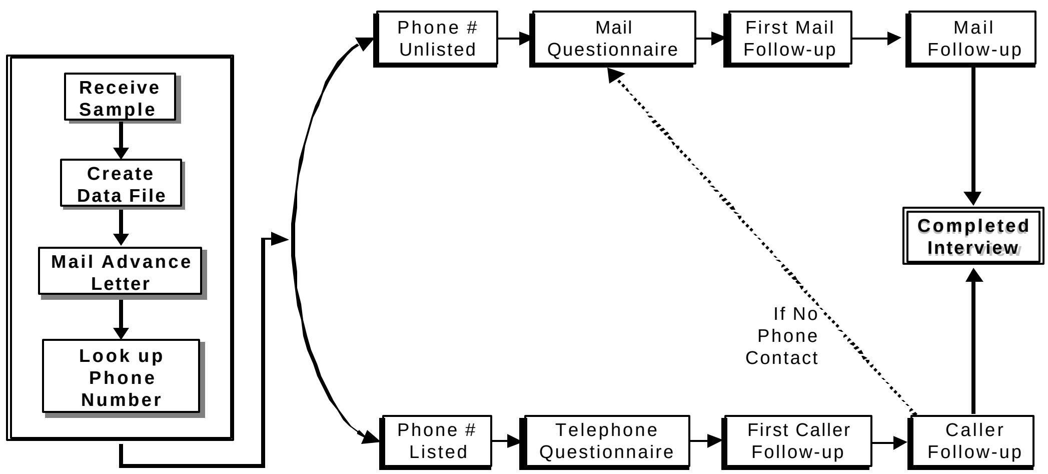
<table>
<tr><th></th><th colspan="3">Self Preparer</th><th colspan="3">Paid Preparer</th><th colspan="3">Software Preparer</th></tr>
<tr><th>Adjusted Gross Income</th><th>Less than $30,000</th><th>$30,000 - $70,000</th><th>Over $70,000</th><th>Less than $30,000</th><th>$30,000 - $70,000</th><th>Over $70,000</th><th>Less than $30,000</th><th>$30,000 - $70,000</th><th>Over $70,000</th></tr>
<tr><td>1040EZ</td><td colspan="3">14,225</td><td colspan="3">1,192</td><td colspan="3" rowspan="3">1,248</td></tr>
<tr><td>1040A</td><td colspan="3">2,828</td><td colspan="3">1,668</td></tr>
<tr><td>1040 & no Schedule A or D</td><td rowspan="3">1,983</td><td colspan="2">4,042</td><td rowspan="3">2,214</td><td colspan="2">6,356</td></tr>
<tr><td>1040 & Schedule A but no D</td><td>1,803</td><td>3,366</td><td>4,019</td><td>1,027</td><td rowspan="2">1,825</td><td colspan="2">10,101</td></tr>
<tr><td>1040 & Schedule D</td><td>10,051</td><td>2,780</td><td>745</td><td>2,288</td><td colspan="2">15,717</td></tr>
</table>

III.2 SURVEY MODE

In developing the survey methodology, several factors were considered in selecting the most appropriate data collection approach. The survey budget, desired response rate, data quality, survey topic, and the population of interest were all factors that were weighed in this decision. While the data needed from taxpayers is difficult to obtain under any data collection design, we believed our mixed mode approach—

telephone and mail—was very effective in achieving the goal response rate and eliciting the information needed to accurately estimate burden. ***Figure III-1*** illustrates the mixed mode process.

FIGURE III-1: MIXED MODE FLOW CHART



The proposed approach enhanced response rates in three ways. First, it relied initially and primarily on telephone surveys, which typically achieve a higher response rate than mail surveys. Second, it provided a means to contact taxpayers for whom no phone numbers were available. Third, it offered an avenue for follow up beyond repeated telephone calls. Taxpayers who could not be contacted by phone were sent a mail survey, resulting in additional responses.

The carefully crafted survey methodology of phone and mail helped to improve recall and consistency in data collection. A related benefit of the initial emphasis on phone surveys was that trained interviewers were able to reduce recall bias by prompting respondents. Carefully crafted mail questionnaires also helped to improve recall. However, the overall quality of mail survey data may still have been marginally lower than the quality of telephone survey data. To ensure that the two data collection modes produce comparable data, both mail and telephone survey instruments were pre-tested, and any systematic differences in responses were explored carefully.[10]

An advance letter was mailed to sampled taxpayers prior to the start of data collection to notify them of the study, explain the importance of defining W&I taxpayer compliance burden, and encourage participation. Thus, the letter notified the target respondents of the study and it was not a surprise when the first call was placed and that the purpose of the call was a legitimate research study. The message was conveyed in such a manner that the respondent believed they were receiving a personal benefit if they participated. An additional letter was included in this mailing that was written on IRS stationary, and signed by the IRS Commissioner.

III.2.1 RESPONSE RATE

In any survey research project, the need to obtain a good response rate must be balanced with the quality of information collected from the respondent. It is important to consider the various response effects associated with conducting survey research. PwC Consulting took into consideration the following to reduce the response effects of the study: the method of survey administration, the number of open versus closed questions contained in the survey instrument, the question order, the length and wording of questions, and the recall ability of the respondents. In order to complete the 6,000 interviews, we selected a sample of 11,086 tax returns. The final response rate was 60.5 percent.

III.2.2 DATA CLEANING

Prior to estimation all survey data were reviewed, edited and cleaned. These measures fall into three broad categories:

- **Checking across questionnaire sections for inconsistencies in related questions.** PwC ensured that survey data were internally consistent and were within reasonable ranges. For example, a respondent could have indicated no use of a paid professional, but went on to provide an out-of-pocket expense figure for a tax professional fee paid.

- **Verification of extraordinary responses.** We did telephone follow-up, where possible, to verify extraordinarily high or low time and out-of-pocket burden responses for accuracy. Respondents who gave burden levels outside of the acceptable ranges were called back to verify responses.

- **Omission or conversion of responses due to failure to follow instructions.** For example, respondents may have provided a range response (e.g., “6-8 hours”) when an integer response was elicited (here, the value was changed to “7”). Some respondents failed to ignore questions due to skip patterns, or provided multiple responses to a single response question (e.g., “check only one”), and these responses were deleted.

[10] Please see Section III.2.2 Data Cleaning for more information.

III.3 DATA COLLECTION METRICS

As stated previously, our goal was to complete 6,000 interviews with W&I taxpayers, and in the end a total of 6,366 interviews were completed. Approximately 60 percent (3,815) of these interviews were completed by telephone.

III.3.1 NUMBER OF RESPONSES BY STRATUM

Our initial survey data contained 6,366 completed interviews. However, after the survey data was matched to Tax Year 1999 RTF data only 5,851 completed interviews were included in the estimation.[11] The distribution of the 5,851 cases is presented in ***Table III-3***. Note that taxpayers can move across our sampling strata from one year to the next—due to changes in AGI, preparation type, or form and schedule type. As a result, the distribution of completes across strata varies depending on the tax year in question.

As we stated previously the overall response rate was 60.5 percent. ***Table III-4*** presents the response rates for each strata as well as the number of completed interviews before and after the review of the Tax Year 1999 RTF data. Strata description, the 1998 W&I tax return population the sample frame and sample size are also presented. ***Table III-4*** presents the distribution of the aggregate sample.

[11] See Section IV.2.1 for a discussion of observations dropped during the merge of survey and RTF data.

TABLE III-3: POPULATION AND SAMPLING FRAME COUNTS FOR W&I TAXPAYERS

Strata	Form and Schedule Type	Preparation Method	Adjusted Gross Income	Tax Year 1998 W&I Population	Sample Frame	Sample Size	Number of all completes by TY1998 strata	Number of final survey data by TY 1999 strata [1]	Response Rate
A	1040EZ	Self	All AGI Categories	14,225,077	1,647	635	282	384	47.0%
B	1040A	Self	All AGI Categories	2,828,255	1,486	521	313	280	63.5%
C	1040 & no Schedules, 1040 & Schedule A, 1040 & Schedule D	Self	Less than $30,000	1,982,669	1,563	516	329	336	67.0%
D	1040 & no Schedules	Self	$30,000 or greater	4,042,206	1,533	526	293	304	57.7%
E	1040 & Schedule A	Self	$30,000 - $70,000	1,802,667	1,467	570	308	256	55.9%
F	1040 & Schedule A	Self	Over $70,000	3,365,958	1,509	549	299	217	56.8%
G	1040 & Schedule D	Self	$30,000 - $70,000	10,050,951	1,538	647	269	303	44.9%
H	1040 & Schedule D	Self	Over $70,000	2,779,610	1,496	718	278	160	43.4%
I	1040EZ	Paid	All AGI Categories	1,192,334	1,151	555	346	215	64.7%
J	1040A	Paid	All AGI Categories	1,668,153	1,335	500	355	457	74.6%
K	1040 & no Schedules, 1040 & Schedule A, 1040 & Schedule D	Paid	Less than $30,000	2,124,155	1,445	512	331	364	66.9%
L	1040 & no Schedules	Paid	$30,000 or greater	6,356,008	1,778	581	284	348	53.2%
M	1040 & Schedule A	Paid	$30,000 - $70,000	4,018,833	1,719	577	308	242	57.4%
N	1040 & Schedule A	Paid	Over $70,000	1,027,334	1,880	500	358	257	75.5%
O	1040 & Schedule D	Paid	$30,000 - $70,000	744,778	2,075	500	378	294	79.4%
P	1040 & Schedule D	Paid	Over $70,000	2,288,217	1,549	500	330	271	67.8%
Q	1040EZ, 1040A, 1040 & no Schedules	Software	All AGI Categories	1,247,745	1,587	500	343	303	69.9%
R	1040 & Schedule A, 1040 & Schedule D	Software	Less than $30,000	1,825,220	1,565	525	329	216	64.6%
S	1040 & Schedule A	Software	$30,000 or greater	10,101,336	1,384	556	300	311	56.3%
T	1040 & Schedule D	Software	$30,000 or greater	15,717,462	1,511	598	303	333	53.3%
			Total	**89,388,968**	**31,218**	**11,086**	**6,366**	**5,851**	**60.5%**

[1] Approximately 500 of the completed interviews were not used in final estimation because the individuals were no longer wage & investment taxpayers or a matching RTF tax return record was not provided. For a further discussion on RTF and the matching of survey data, please see Section IV.2.1.

TABLE III-4: AGGREGATE SAMPLE DISPOSITION

COMPLETED INTERVIEWS	
❖ Completed by phone	3,815
❖ Completed by mail (never contacted by phone)	1,397
❖ Sent questionnaire and completed by mail	1,154
Subtotal Completed Interviews	**6,366**
INCOMPLETE	
❖ Unresolved from Mail	1,688
❖ Refusal	874
❖ Phone Number Not Available	673
❖ Maximum Number of Call Attempts Reached (20)	499
❖ Data Collection Ended Before Response	245
❖ Respondent Injured / Ill / Elderly	60
❖ Requested Mail Version but did not Complete	50
❖ Respondent Traveling/ Out of Country	27
❖ Insufficient Completion of Mail	20
❖ Wrong Respondent from Mail	13
❖ Respondent Terminated Interview	13
❖ Respondent Requested Phone interview	1
Subtotal Incomplete	**4,163**
OUT-OF-FRAME	
❖ Screened out, e.g., not W&I; did not file, etc.	255
❖ Incorrect Address for mail version	144
❖ Language problem	93
❖ Respondent Deceased	47
❖ Respondent no longer at this address	11
❖ Other	7
Subtotal Out-of-Frame	**557**
Grand Total	**11,086**
Response Rate (complete/(complete + incomplete))	**60.5%**

IV. ESTIMATION

Of all the tasks involved in building a microsimulation model, perhaps the most challenging is to develop the algorithms that forecast taxpayer behavior and burden. A number of factors contribute to the complexity of this task. There are several outcomes of interest, including filing outcomes, compliances methods, and time and money burden. Each outcome is influenced by a wide range of factors, some of which are unobservable or difficult to quantify (e.g., taxpayer aptitude). The algorithms must reflect enough of these factors to provide accurate forecasts of taxpayer behavior, but not so many that the model becomes intractable—and collectively, the factors included in each algorithm must provide model users with simulation levers that accommodate a wide variety of what-if scenarios.

In this section, we discuss the process of specifying and estimating algorithms for the burden model. Section IV.1 reviews the theoretical underpinnings of the model and motivates the data needs for the estimation process. Section IV.2 describes the development of an estimation data file. Section IV.3 documents our estimation methodology, presents the results of our estimation, and reviews a number of diagnostics on model fit and reliability.

IV.1 THEORETICAL OVERVIEW

Almost every empirical investigation of economic behavior is based on some underlying logical structure that describes the behavior of the agents in the system, and is the basic framework for analysis. This logical structure is put forth in the form of equations that describe the behavior of economic and related variables. The model may consist of a single equation or as a system involving many interrelated equations. The ultimate goal is to empirically estimate the theoretical relationship, and draw conclusions from the findings.

In Section II, we identified three major components of the burden model aimed at simulating taxpayer behavior—a tax engine to simulate filing outcomes, a decision module to simulate compliance methods, and a burden module to simulate the level of time and money burden. To guide the development of these components, we first sought to understand the theoretical relationship between each outcome and its determinants.

IV.1.1 FILING OUTCOMES

Filing outcomes are the values reported on each line item of a tax return. These outcomes are simulated in the Tax Engine using a series of algorithms that, in effect, represent the rules of the federal income tax system for a given tax year. While complex and voluminous, these algorithms are generally deterministic. For example, in TY99, a taxpayer whose taxable income exceeds $50,000 must file Form 1040, and a taxpayer who has taxable Social Security income must report that income on his return. This deterministic quality becomes nearly universal if one assumes that all taxpayers seek a tax-minimizing outcome. Under this assumption, a taxpayer will choose to itemize if his itemized deductions are greater than his standard deduction and will claim any tax credit for which he is eligible.

In reality, not all taxpayers achieve minimum-tax outcomes. Some are unaware of their eligibility for credits and deductions, while others may choose to pay additional tax in order to reduce their compliance burden (e.g., claiming the standard deduction to avoid calculating itemized deductions). Recognizing this fact, we made two simplifying assumptions that allow us to simulate filing outcomes in a deterministic manner: (1) taxpayers generally choose filing outcomes that minimize their tax liability, and (2) taxpayers who display non-tax-minimizing behavior in the base-year data will continue this behavior in all subsequent scenarios.[12]

IV.1.2 COMPLIANCE METHODS

As mentioned above, the importance of tax compliance methods as determinants of burden led us to include a separate component of the burden model to predict taxpayer decisions with respect to these methods. Through discussions with the working group and other stakeholders, we identified four types of compliance methods of particular interest: (1) preparation method, (2) submission method, (3) use of IRS services, and (4) methods used to gather tax materials (e.g., forms and publications, tax manuals). Below, we discuss the factors that influence taxpayer decisions regarding each compliance method.

IV.1.2.1 PREPARATION METHOD

Prior research indicates that most taxpayers exhibit inertial behavior with respect to preparation method—choosing the same method from year to year. When a change does occur, it is often the result of a life event such as the birth of a child or the purchase of a house. In many cases, these life events have tax implications, which result in an increase in the complexity of the taxpayer's return. For example, first time homebuyers often find that it is advantageous to itemize deductions, and therefore must file a Schedule A for the first time. These dynamics lead us to believe that the best theoretical model for

[12] Because we assume deterministic selection of filing outcomes, the Tax Engine's algorithms are based on rules rather than a statistical model. For a more detailed discussion of the Tax Engine's algorithms, see the report, Tax Engine Algorithms (May 2001).

preparation method is one that predicts changes from an initial state as a function of changes in life events and key filing outcomes.

IV.1.2.2 SUBMISSION METHOD

As with preparation method, taxpayers exhibit inertial behavior with respect to submission method—but this inertia is tempered by several competing influences. First, the choice of preparation method often dictates or influences the choice of submission method. Taxpayers who prepare their return by hand have few options but to submit their return by mail, while software preparers are encouraged to file electronically. Second, taxpayers who expect a refund are much more likely to submit their return electronically than are those who have a balance due. This is true both because electronic filing results in a faster refund, and because most taxpayers choose to pay their balance due by mail. Third, IRS is promoting the use of electronic filing through a variety of initiatives, including the self-select PIN program, expanded Fed-State Telefile, and new web-based payment options. Again, we believe that the best theoretical model for submission method is one that predicts changes from an initial state as a function of changes in these other factors.

IV.1.2.3 USE OF IRS SERVICES

The IRS offers a number of services to assist taxpayers in the compliance process and, by extension, reduce taxpayer burden. Among these services are Volunteer Income Tax Assistance (VITA), Tax Counseling for the Elderly (TCE), and TeleTax. The factors that lead taxpayers to use these services are not well understood, but data collected in this study indicate that service use is correlated with preparation method (self-preparers are the most frequent users of IRS services). Taxpayer demographics may also be related to the propensity to use some IRS services, since the services themselves are geared towards specific subgroups.

IRS's first Annual Report on Tax Law Complexity suggests that certain complex filing requirements are also triggers for service use. That report identified the ten tax issues or questions most frequently asked by taxpayers through a variety of channels, including toll-free tax assistance, TeleTax, and internet e-mail. Many of these questions that prompt taxpayers to use IRS services relate to three broad areas of complexity identified in the report: Filing Definitions, Estimated Tax, and Alternative Minimum Tax.

IV.1.2.4 GATHERING TAX MATERIALS

Each year, taxpayers gather a variety of tax materials to help them understand and comply with tax laws. Many taxpayers retrieve forms and publications, which the IRS makes available through several channels, including libraries, post offices, the Digital Daily Web site, and the Tax Fax service. Others purchase tax manuals or attend seminars that help them with tax planning.

A number of factors may influence a taxpayer's decision to gather materials. Preparation method is a major factor—taxpayers who rely on a paid professional are unlikely to need additional materials, and those who use tax software have access to most forms and publications through the software package. Self-preparers are much more likely to gather tax forms or publications, particularly if the mail package they receive from IRS is missing a relevant form, or if a change in their tax situation forces them to deal with unfamiliar tax rules. Intuition also suggests that taxpayers dealing with unfamiliar situations are more likely to purchase tax manuals or attend seminars.

IV.1.3 TAXPAYER BURDEN

The level of taxpayer burden is the central focus of the model. As such, it is appropriate that burden be measured and simulated at a relatively fine level of detail. Indeed, the taxpayer survey provides 14 separate measures of burden—time burden and money burden for each of seven activity categories. There are two main reasons for measuring burden by activity category. First, tying the burden to specific activities helps survey respondents recall all of the time and money they spent on tax compliance, thereby increasing measurement accuracy. Second, it facilitates the development of explanatory models, as different types of burden may be influenced by different factors. For example, while return complexity may strongly influence the time it takes to complete a tax return, it may have no impact on the cost of submitting the return.

Conceptually, we can think of the level of taxpayer burden as a function of three groups of factors:

- **Taxpayer Activities.** Since taxpayer activities are generally not observed directly, we must rely on the contents of the tax return (i.e., filing outcomes) to infer the underlying activities. This approach produces sound results for some activities (e.g., form completion), but is more tenuous for others (e.g., tax planning). To fill in important gaps, some activity-related questions are included in the taxpayer questionnaire.

- **Compliance Methods.** As discussed in Section IV.1.2, compliance methods are the tools, services, and techniques that a taxpayer uses to achieve his or her filing outcomes. Preparation method is by far the most important of these methods, affecting nearly all aspects of the compliance process.

- **Taxpayer Ability.** This category includes a variety of demographic characteristics that describe proficiency with regard to tax compliance activities. Examples could include education, familiarity with specific tax rules, computer literacy, and language proficiency.

We expect that these factors relate to burden in a non-linear fashion. For example, taxpayers who are familiar with a given tax form will generally spend substantially less time on that form than would a first-time filer. However, increasing familiarity with the form probably yields diminishing time-savings. The relationship between burden and its determinants is also expected to be interactive. For example, taxpayer experience may be extremely important among self-preparers but irrelevant among taxpayers that use a paid professional.

Drawing on the theoretical relationships described above, we identified three principals that guided our estimation of taxpayer burden. First, recognizing that different types of burden are influenced by different factors, we opted for a modular estimation approach that could generate separate estimates of time and money burden for each activity category. Second, we decided that the estimation technique selected for each activity category should depend on several criteria, including the average level of burden, the dispersion of that burden, and the importance of the activity from a simulation perspective. Third, we anticipated that data reduction techniques would be needed to avoid problems of multicollinearity, due to the large number of explanatory factors in some econometric equations.

IV.2 Creating an Estimation Data File

In order to empirically test the theoretical relationships described above, we needed a data file with information on each outcome variable, as well as a wide range of explanatory variables. To create this data file, we merged records from three data sources—the W&I taxpayer survey, the TY99 Returns Transaction File (RTF), and the TY98 RTF.[13] ***Table IV-1*** illustrates the types of information contained on each data file.

[13] All returns processed in a given filing year (i.e., calendar year) are included in the RTF for that filing year, regardless of the tax year of the return. Thus, while the 2000 RTF consists predominantly of TY99 returns, it also includes returns from earlier tax years. When we refer to the TY99 RTF, we are referring to all TY99 returns in the 2000 RTF.

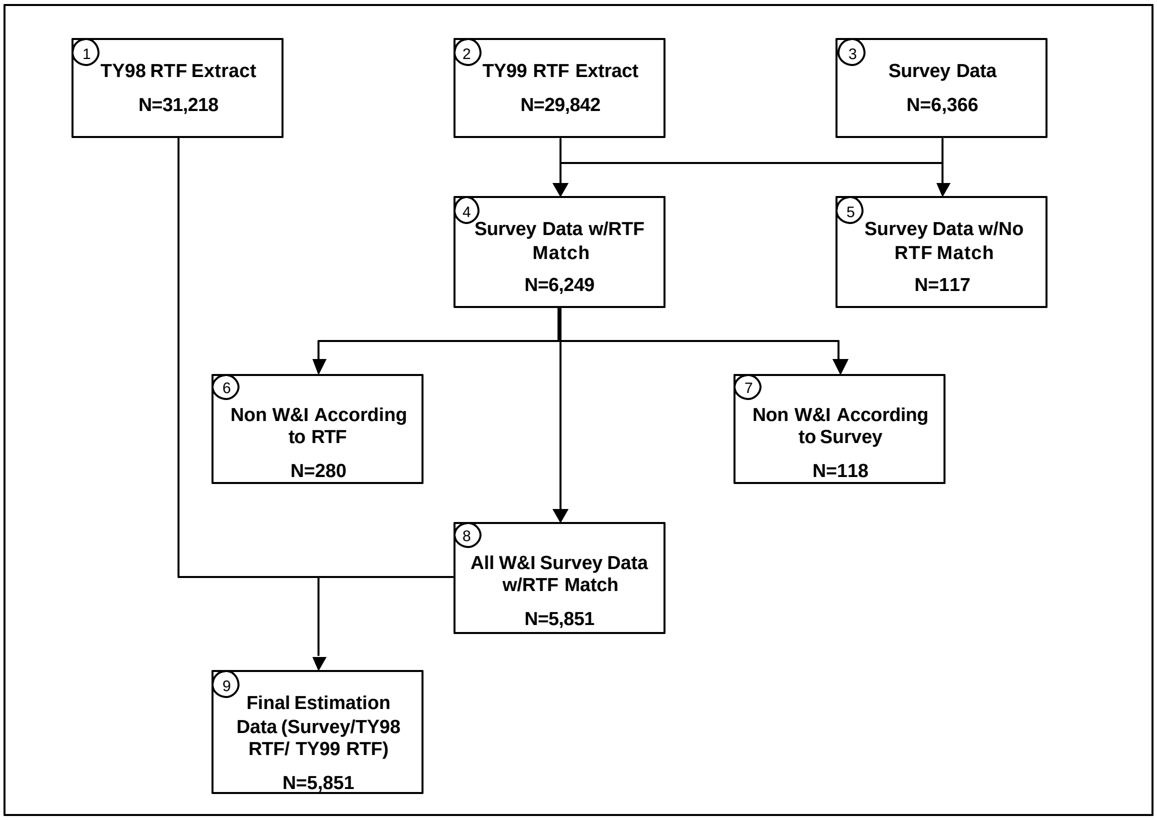
TABLE IV-1: INFORMATION PROVIDED BY EACH ESTIMATION DATA SOURCE

Data Element	Taxpayer Survey	TY99 RTF	TY98 RTF
OUTCOME VARIABLES			
❖ Preparation Method	✓	✓	✓
❖ Submission Method	✓	✓	✓
❖ Use of IRS Services	✓		
❖ Gathering Tax Materials	✓		
❖ Time and Money Burden	✓		
TAXPAYER CHARACTERISTICS			
❖ Computer/Internet Access	✓		
❖ Education	✓		
❖ Experience	✓		
❖ Income and Expenses	✓	✓	✓
❖ Location of Residence	✓	✓	✓
❖ Marital Status / Dependents	✓	✓	✓
TAXPAYER ACTIVITIES			
❖ Filing Outcomes		✓	✓
❖ Forms not Submitted	✓		
❖ Record Keeping Activities	✓		
❖ Tax Planning Activities	✓		
❖ Working with a Paid Professional	✓		

IV.2.1 MERGING SURVEY DATA AND RTF DATA

The survey and RTF data files were merged using a unique, record-level identifier variable, which was created by the IRS as a proxy for the primary taxpayer's SSN. ***Figure IV-1***, below, illustrates the process and results of merging the three data files.

FIGURE IV-1: ESTIMATION FILE DEVELOPMENT PROCESS



For a variety of reasons, not all of the 6,366 survey respondents (box 3) are retained in the final estimation data file. In 117 cases (1.8 percent), IRS was not able to provide a TY99 RTF record for the taxpayer (box 5). This may indicate that the taxpayer did not file a return in TY99—although all survey respondents reported filing a return—or that the taxpayer's return could not be located in the RTF.[14] In another 398 cases (6.3 percent), the respondent appeared to be a Self-Employed taxpayer, either based on RTF data (box 6) or survey data (box 7).[15] After excluding these records, 5,851 Wage and Investment records remained in the final estimation data file (box 9).

IV.2.2 INCONSISTENCIES BETWEEN SURVEY AND RTF DATA

As shown in ***Figure IV-1***, there were 118 cases where survey respondents reported filing one or more Self-Employed forms in TY99 but the RTF data did not support this claim. This mismatch is an example

[14] The absence of these records in the RTF may indicate that some returns had not been posted as of production cycle 200052 (end of calendar year 2000), or that the primary SSN selected from the TY98 RTF was not present as a primary or secondary SSN in the TY99 RTF.

[15] The taxpayer questionnaire included a series of screener questions to establish whether the taxpayer filed any Self-Employed tax forms (Schedules, C, E, or F, or Form 2106). Telephone interviews were terminated if a respondent reported filing any of these forms, but self-administered mail questionnaires were continued.

of the inconsistencies that can arise when comparable outcomes are observed through two data sources. The three inconsistencies most commonly observed between survey data and RTF data are inconsistencies in preparation method, submission method, and primary tax form filed.[16]

Comparing survey and RTF data in terms of preparation method is complicated by the fact that survey respondents were allowed to report multiple preparation methods. To resolve this complication, each survey respondent was assigned a single method based on the following hierarchy: (1) paid preparer, (2) software preparer, (3) self preparer.[17] In most cases, this hierarchically assigned preparation method is consistent with the preparation method recorded in the RTF. Between 82 and 88 percent of taxpayers assigned to each survey-based category have a matching RTF preparation method. Similarly, most taxpayers in each RTF category have a matching survey-based preparation method. One noteworthy deviation from this trend is that, of the 14.0 million (weighted) taxpayers that use tax software according to the RTF, only 8.5 million (60.4 percent) are classified as software users through the survey-based hierarchy.

Survey and RTF data are generally consistent in terms of return submission method as well (i.e., paper, TeleFile, other electronic). Between 72 and 89 percent of taxpayers in each survey-based category have a matching RTF submission method. However, RTF records are substantially more likely to indicate that a return was filed electronically. Of the 6.7 million taxpayers that TeleFile according to RTF, only 2.2 million (32.0 percent) report TeleFiling in the survey data. Similarly, of the 26.9 million taxpayers that use some other electronic submission method, only 11.5 million (42.6 percent) report the same method in the survey data. We suspect that this discrepancy is caused in part by respondent error, either due to uncertainty over how the return was filed (e.g., a taxpayer whose paid professional filed the return electronically) or uncertainty over what constitutes an electronically filed return (e.g., a taxpayer who submits his return electronically, but mails in his W-2s and/or balance due).

Finally, we observed some inconsistency between survey and RTF data in terms of the primary tax form filed. The degree of this inconsistency is difficult to quantify, primarily because mail respondents were not asked to report their primary tax form. Nevertheless, of those CATI respondents who reported filing

[16] For a more detailed discussion of inconsistencies between survey and RTF data, see the "Data File Development" section of the document Final Estimation Summary (June, 2001).

[17] For example, a taxpayer who used all three preparation methods would be assigned to the paid-preparer method; a taxpayer who self-prepared and used tax software would be assigned to the software preparer method. For purposes of this hierarchical assignment, receiving assistance from a friend, relative, or IRS representative, was treated as equivalent to self preparation.

Form 1040, 1040A, or 1040EZ, approximately one-third reported a primary form that was inconsistent with the that observed on the RTF.

While it is not possible to determine which data source is correct in each instance of an inconsistency, one can devise general rules that help to resolve these inconsistencies. We relied on three such rules. In cases where the survey and RTF files differ in terms of the forms and schedules filed, we rely on the RTF data—reflecting our assumption that taxpayers may be uncertain about the specific forms they filed.[18] In cases where the survey and RTF data differ in terms of submission method, we also rely on the RTF data, due to apparent underreporting of electronic filing in the taxpayer survey. Finally, in cases where the two data files differ in terms of preparation method, we rely on the survey response. The reason for this decision is twofold. First, we believe that taxpayers can accurately recall this fundamental tax compliance method, and second, given the strong relationship between preparation method and taxpayer burden, it is essential that the preparation method reported on the estimation data file be consistent with the burden levels reported on the survey.

IV.2.3 Variable Creation

While the merged survey/RTF data provided much of the information needed to support model estimation, hundreds of variables had to be constructed to translate this raw data into useful predictor variables. In this section, we describe three types of constructed variables that were instrumental in developing explanatory models—indicator variables, change variables, and attribute variables.

IV.2.3.1 Indicator Variables

RTF data provides a detailed account of the dollar amounts reported on a tax return, but the relationship between these dollar amounts and taxpayer burden is tenuous. A high-income taxpayer whose only source of income is wages may incur substantially less burden than a low-income taxpayer whose income is derived from several different sources. Similarly, the burden associated with income from capital gains may have more to do with the number and complexity of transactions than with the dollar amount of the gain. Speaking more generally, it is unclear which is a more robust predictor of taxpayer burden and other key outcomes—the dollar amounts associated with particular filing outcomes or indicators of those filing outcomes.

[18] The one exception to this rule is that taxpayers who reported filing a self-employed form were excluded from the estimation sample, even if the RTF data indicated that they did not file a self-employed form.

In order to test both relationships, we developed hundreds of indicator variables that reflect the forms and line-items completed in a return.[19] Generally, we required a non-zero value for a line-item as evidence that the line-item was completed. In some cases, however, we were able to infer that a taxpayer must have worked on a line item even if a value of 0 was reported or if the RTF file was missing data for that line-item. For example, if a taxpayer's itemized deductions on Schedule A are limited, we can infer that the taxpayer completed the itemized deduction worksheet for line Schedule A, Line 28.

IV.2.3.2 CHANGE VARIABLES

As mentioned above, taxpayers tend to display inertial behavior with respect to preparation and submission methods, leading us to favor a theoretical model that predicts changes from an initial state as a function of changes in life events and key filing outcomes. To estimate such a model, we needed to create a number of variables describing changes in compliance methods, as well as changes in filing outcomes (e.g., filing a form that was not filed in the prior tax year) and taxpayer characteristics (e.g., birth of a child). In addition to these binary change variables, we created many continuous change variables (e.g., change in tax liability, change in non-wage income), which were hypothesized to be triggers for other burden activity categories (e.g., tax planning).

IV.2.3.3 ATTRIBUTE VARIABLES

Collectively, the indicator variables described above offer a proxy for the volume and complexity of compliance activities encountered by a taxpayer. Unfortunately, the sheer quantity of these indicators makes them difficult to use in an estimation model. Moreover, the fact that they are indicators only of *current* filing outcomes limits their value when trying to simulate the impact of future filing outcomes. To overcome these two weaknesses, a new class of variables (attribute variables) were created with two primary objectives: (1) to quantify the volume and complexity of all existing filing outcomes using a smaller number of variables, and (2) to measure volume and complexity in a way that allows new filing outcomes to be measured on an identical scale.[20]

Attributes are characteristics of tax rules or requirements that allow us to infer a taxpayer's activities based on his or her filing outcomes. In establishing a set of attributes to measure, several criteria had to be met. First, the set of attributes should be comprehensive—describing both a wide range of factors that

[19] These variables take a value of 0 if the form/line was not completed and a value of 1 if the form/line was completed.

[20] For a more detailed discussion of the attribute methodology, see Section III of the report, Specifications for the Interactive Decision Support Tool (September 14, 2001), and the "Data File Development" Section of the Final Estimation Summary (June 2001).

influence burden (e.g., activity volume, complexity, ambiguity), and a wide range of tax compliance activities (e.g., form completion, record keeping, tax planning). Second, each attribute should be objectively defined, so that the attributes associated with a filing outcome are not subject to interpretation. Third, the attributes should be easy to measure, both for existing filing outcomes and for new filing outcomes.

The attribute framework we have developed attempts to balance these criteria by using three distinct types of attributes—source attributes, operation attributes, and complexity attributes. This attribute framework was designed based on the notion that tax compliance burden is primarily a function of three things: (1) the information the taxpayer has to provide, (2) the operations the taxpayer performs on that information, and (3) the difficulty of gathering the information and performing operations. Source attributes describe the information source for a given filing outcome—such as an information return or a worksheet. Operation attributes describe the operations performed in order to realize a filing outcome—such as calculations, comparisons, or consulting a lookup table. Complexity attributes describe factors that influence the difficulty of performing the aforementioned activities—such as exceptions to the standard tax rules for certain individuals or certain income types.

In all, the attribute framework encompasses 20 distinct attributes—5 source attributes, 8 operation attributes, and 7 complexity attributes. To mitigate problems of multicollinearity in the estimation, we further reduced these 20 attribute variables into a unidimensional index through the use of principal components analysis. The resulting "attribute index" is the single value that captures the greatest amount of information in the underlying attributes. This index is computed by multiplying each standardized attribute variable by the corresponding principal component coefficient—then summing across the 20 attributes.[21]

IV.3 ESTIMATION METHODOLOGY AND RESULTS

IV.3.1 DECISION MODULE

The decision module uses logistic regression (logit) equations to simulate two primary decisions for each taxpayer—the choice of preparation method and the choice of submission method.[22] For a subset of

---

[21] Each standardized attribute variable is computed as the number of standard deviations away from the mean value for the attribute.

[22] The logit is one of the most widely used qualitative choice models. The popularity of the logit is due to the fact that the formula for logit choice probabilities is readily interpretable and fairly easy to estimate with most statistical software packages. Logistic regression models are typically used to estimate binary choices of individuals and multinomial logit models are used to model relationships between a polytomous response variable and a set of regressor variables.

taxpayers, the decision module also simulates the decision to TeleFile. In the following section, we discuss the process of specifying and estimating the equations for choice of preparation method. Section IV.3.1.2 reviews the estimation procedures used to model the choice of submission method.

IV.3.1.1 PREPARATION METHOD

In an effort to model the impact of various factors on the choice of preparation method, we chose to estimate a multinomial logit model. These response models are typically classified in two distinct types—ordered and unordered—depending on whether the response variable has an ordered or unordered structure. An example of an ordered response variable might be severity of a medical condition, with response categories of none, mild, and severe. This variable is ordered in the sense that “none” is less than “mild” is less than “severe”. In the case of preparation method, the response variable has an unordered structure. Taxpayers choose one of three preparation methods—paid professional, self-prep without software, or self-prep with software—but there is no inherent order to these methods.

The unordered choice of taxpayer preparation method can be motivated by a random utility model. For the n^{th} taxpayer, faced with three choices of preparation method, suppose that the utility of choice j is:

$$U_{nj} = bx_{nj} + e_{nj}$$

If we observe that a taxpayer chooses a particular preparation method (e.g., self-preparation without software), we can infer that this choice yields the maximum utility among the three alternatives. Therefore, a statistical model is guided by the probability that choice j is made, which is:

$\Pr(U_{nj} > U_{nk})$, for all other k choices of preparation method.

This statistical model of preparation method is made operational through an assumption regarding the distribution for the disturbance term. As noted earlier, the most common model of this type is the multinomial logit model, which assumes a logistic distribution of the disturbance. Consider the choice of preparation method. The alternatives facing an individual taxpayer include (1) using a paid professional, (2) self-preparation without software, and (3) self-preparation with software. The multinomial logit has a particularly simple mathematical representation for the probability that a particular choice is made:

$$\Pr(Y = j) = \frac{e^{b_j x_n}}{1 + \sum_{k=1}^{J} e^{b_k x_n}}$$

The estimated equations provide us with a set of probabilities for the three preparation methods for a taxpayer with characteristics x_n. We explored many specifications for the variables included in x_n and in the final model we chose the following explanatory factors:

(1) Preparation method in the prior year

(2) Change in the primary tax form filed (1040, 1040A, or 1040EZ), and

(3) Net change in the number of non-primary forms filed

This parsimonious specification has several advantages. First, all of the variables in the equations were found to be highly significant in the estimated multinomial logit. Second, this specification preserves much of the inertial behavior exhibited in the data. Third, this specification ensures that the simulated choice of preparation method will be responsive to a number of policy levers available through the software, including tax law parameters, tax law structure, and taxpayer demographics.

Table A.1 (Appendix A) presents the estimated probabilities for each of the possible combinations of the independent variables. As an example of how the estimated model for preparation method can be interpreted, consider an individual with the following characteristics: (1) the taxpayer used a paid professional in the prior year, (2) the taxpayer transitioned from a 1040 in 1998 to a 1040EZ in 1999, and (3) the taxpayer had a net decrease in non-primary forms.[23] The estimated model predicts that this individual has a 47 percent chance of remaining with a paid professional, a 49 percent chance of switching to self preparation without software, and a 4 percent chance of switching to self preparation with software.

The model we estimated has much econometric and intuitive appeal, and satisfied our objectives of having a model that was grounded in theory, statistically significant, and sensitive to policy levers that could be incorporated into the software. In the example above, the model predicts that individuals with a reduction in the complexity of their return will have a greater probability of self-preparing, but will continue to be influenced by inertial behavior.

IV.3.1.2 SUBMISSION METHOD

As in the estimation of preparation method (which relied on a multinomial logit model) submission was estimated using a binary logit model. The submission method outcome can take one of two possible

[23] Moving from 1040 to 1040EZ might have also meant the elimination of Schedule B, since 1040EZ filers should report less than $400 in interest income.

values—paper or electronic—for software preparers and those who use a paid professional. For self-preparers who do not use software, the electronic submission option is only available for those who are eligible to TeleFile. These restrictions are reflected in our estimation.

The strongest determinant of submission method in our model is submission method in the prior year—again demonstrating that taxpayers exhibit strong inertial behavior. The only other strong predictors of submission method for W&I taxpayers were preparation method and refund status. For example, taxpayers who self-prepared using software in 1999 and submitted electronically in 1998 were much more likely to submit electronically in 1999 if they had a refund due (91 percent) than if they owed tax (62 percent). Tables A.2 through A.4 (Appendix A) contain all of the estimated probabilities for transitioning into and out of electronic filing, including TeleFile.

IV.3.2 BURDEN MODULE

The purpose of the Burden Module is to predict the incidence and level of burden across all of the burden activity categories. These outcomes are predicted in two steps. First, we predict the likelihood that each taxpayer experiences time and/or money burden in each activity category, then we predict the level of burden for those taxpayers that incurred non-zero burden.

A number of candidate models were tested in the estimation of taxpayer burden, including logit, probit, Tobit, and segmentation models. Ultimately, we used segmentation models to predict the incidence of burden—and a combination of segmentation models and econometric equations to predict the level of burden. The decision on which type of model to use for burden levels was guided by several factors, as discussed below.

IV.3.2.1 BURDEN ESTIMATION FRAMEWORK

The models used to predict the level of time and money burden included econometric equations and segmentation rules. The estimation approach employed was influenced by five primary factors:

(1) **The percent of taxpayers that engaged in a particular activity.** Activities that affected a small percentage of taxpayers were generally given lower priority than those that affected a large percentage of taxpayers.

(2) **The mean level of burden within an activity category.** Related to the point above, activity categories that encompassed a large proportion of total time or total money were generally considered to have greater importance. For example, the mean amount spent on a paid professional by individuals who used a paid professional to prepare their tax return was approximately $113, the mean amount spent on form submission was

approximately $4. Given the difference between the these two, we placed a higher priority on developing models for paid professional money.

(3) **The distribution of burden within an activity category.** It is inherently difficult to develop models for (burden) variables with little variation. Since the data are clustered around the mean, there is little variance "left to explain" outside of the mean of the variable. An example of this is the distribution of form submission time. Form submission time had little variation, even across preparation methods.

(4) **The availability of data that could be used in an explanatory model.** In many cases theory suggested several potential variables that could be included in a model of burden for a particular activity category. Due to data limitations, those variables (or reliable proxies thereof) were not present at the time of estimation. For example, knowledge of or familiarity with tax law may have been an important explanatory variable for those taxpayers that self prepared without the use of software. However, this variable was not present on the estimation data file—and while education may be seen as a proxy, complicating factors arise, since there may only be weak association between education level and understanding of tax law.

(5) **The importance of the activity category for running "what-if" scenarios.** Using IRS services and gathering tax materials were given special consideration from the standpoint of model functionality. These Tax Administration policy variables were considered of sufficient import that econometric equations were estimated to ensure the availability of certain policy levers in the simulation model.

The five criteria listed above are summarized in Table B.1 (time burden) and B.2 (money burden) in Appendix B. These matrices present important metrics that guided our decision to favor one type of model over another. Early examination of the data revealed clear distinctions in the five criteria across preparation methods—thus the table is further divided to illustrate these differences. For each activity category, and by each preparation method, five summary metrics are listed: (1) Percent Impacted, (2) Mean Burden, (3) Median Burden, (4) Standard Deviation, (5) Standard Error. Also listed in each cell is our overall assessment of the cell's importance (high, medium, or low), and the type of model ultimately used to estimate burden (segmentation model or econometric equation).[24]

[24] An outlier was defined as an observation with a burden level greater than five standard deviations from the mean for the burden activity category, segmented by preparation method and primary form filed. For example, Paid Professional 1040 filers had mean recordkeeping time of 8.77 hours and the standard deviation was 18.14. Observations within this category were identified as outliers if the value of recordkeeping time was greater than 99.47 hours (five standard deviations above the mean). This method identified 14 records as outliers for recordkeeping within Paid Professional 1040 filers. For a more detailed exposition of outliers, please refer to *Phase B: Wage and Investment Taxpayer Burden Study Group, Final Estimation Summary*, June 2001.

IV.3.2.2 INCIDENCE OF BURDEN

The objectives of estimating burden incidence across activity categories include: (1) to determine, at the micro level, which taxpayers will have a high propensity to engage in various compliance activities, and (2) to provide useful policy levers related to the incidence of burden for various subpopulations. To accomplish this, we relied on segmentation models.

Similar to the methodology used to estimate submission method (i.e., estimating logit equations), these segmentation models make use of a small number of categorical explanatory variables. Observations are grouped into segments, and choice probabilities vary only over the segments, not over individuals within a segment. The choice probability for each segment is calculated as the proportion of taxpayers in that segment that incurred burden:

$$P_s = \frac{\sum_{i=1}^{N_s} b_i w_{is}}{\sum_{i=1}^{N_s} w_{is}}$$

where b_i is a binary variable that equals 1 if the taxpayer incurred the burden of a particular activity category, and 0 if they did not incur the burden. The w_{is} is the weight associated with the i^{th} observation in segment s. N_s is the total number of taxpayers that are classified in segment s.

In order to forecast the incidence of burden, the set of explanatory factors (segmenting variables) cannot logically be changed, however, individuals may move in and out of particular segments. For example, preparation method is used as an explanatory variable in the segmentation model for Gathering Tax Materials Time. Therefore, if a taxpayer's preparation method changes, so too will his likelihood of incurring burden related to gathering tax materials.[25]

In addition to the time and money burden activity categories, we estimated segmentation models for the incidence of selected behavioral outcomes that relate to tax system administration. Specifically, we estimated four segmentation models related to the use of IRS Services: (1) visiting walk-in sites for tax assistance, (2) calling Tele-Tax to hear messages about the law, (3) calling the IRS Toll Free Tax Assistance line, (4) using IRS website for reasons other than downloading forms, and three models related

[25] For detail on the segmentation rules, please refer to *Phase B: Wage and Investment Taxpayer Burden Study Group, Final Estimation Summary*, June 2001.

to methods that taxpayers use to gather tax materials: (1) visiting the library or post office, (2) ordering or downloading forms or publications from the IRS website, (3) obtaining books or guides for tax purposes.

IV.3.2.3 LEVEL OF BURDEN – ECONOMETRIC EQUATIONS

Using the criteria described in Section IV.3.2.1, we identified six high-priority burden categories: (1) money spent on a paid professional by taxpayers that used a paid professional to prepare their return, (2) recordkeeping time, (3) form completion time, (4) using IRS services time, (5) gathering tax materials time, and (6) tax planning time. In this section, we describe the methods used to estimate econometric equations for these variables.

IV.3.2.4 ESTIMATION RESULTS

To specify each econometric model, we started with a set of theoretical models relating time and money burden to explanatory variables, including taxpayer demographics, filing outcomes, and burden attributes. The empirical models we ultimately selected are an attempt to operationalize these theoretical models within constraints imposed by the data. In one of the econometric equations (paid professional money), the specification we chose was linear in both the variables and parameters. All other equations are estimated using a non-linear (logarithmic) transformation on the dependent variables, but are linear in the parameters.

To predict the amount of money spent on a paid professional, we specified a qualitative variable model—sometimes referred to as a “dummy variable” model. We hypothesized that the amount of money spent on a paid professional was a linear function of the forms that the paid professional actually prepared. To test this hypothesis, we defined several dummy variables representing common forms and schedules, such that:[26]

$$D = \begin{cases} 1 \text{ if the form was filed} \\ 0 \text{ if the form was not filed} \end{cases}$$

The coefficients from this equation can be interpreted as the amount of money spent by the taxpayer to have a particular form completed by a paid professional. For example, the coefficient for the Intercept term in Table C. (Appendix C) indicates that the base amount paid by a 1040, 1040A, or 1040EZ filer is approximately $60. Coefficients for other dummy variables in this equation indicate that filing Schedule

A adds about $30 to the cost of a return, and Schedule B adds about $13. In an attempt to make the model responsive to new forms that IRS may introduce in the future, we constructed a variable to capture the number of "other" forms that made up a taxpayers return. Overall, the equation for paid professional money has reasonably good fit, with an R^2 of .189, and the F-statistic of 43.249 indicates strong explanatory power.

All of the econometric equations for time burden were estimated utilizing a natural log transformation in the dependent variable. This type of specification is often referred to as a log-linear or semi-log model. The natural log transformation was chosen because of several desirable properties, including: (1) it ensures that predicted burden levels will always be positive, (2) it allows for a non-linear relationship between time burden and the various independent variables, and (3) its coefficients are relatively easy to interpret (i.e., coefficients for dummy variables can be interpreted as multipliers and coefficients for continuous variables are easily converted to elasticities).

In developing econometric models for time burden, we generally estimated three equations for each activity category—one for each preparation method. This segmentation is illustrated by the three specifications presented below, for form completion time. Segmenting the sample in this way reflected our assumption that the same explanatory variables may affect different taxpayers differently. For example, mathematical calculations may be burdensome for taxpayers who self-prepare without software, but trivial for taxpayers who use software or a paid professional. This assumption was confirmed through statistical tests that verified differences in the value of the same coefficient across the three equations.

Paid Preparers

$$\ln(time)_i = b_0 + b_1 OT_i + b_2 LSM_i + b_3 MFJ_i + b_4 MFS_i + b_5 AI_i + b_6 TF_i + b_7 NS_i + b_8 MPM_i + e_i$$

$$(i = 1,2,\ldots,I)$$

Self Preparers without Software

$$\ln(time)_j = b_0 + b_1 OT_j + b_2 LSM_j + b_3 MFJ_j + b_4 MFS_j + b_5 AI_j + b_6 TF_j + b_7 NS_j + b_8 MPM_j + e_j$$

$$(j = 1,2,\ldots,J)$$

[26] The forms for which dummy variables were created include: Schedules A, B (or 1), D, EIC, and Forms 2441 (Schedule 2), 6251, 1116, 2210, 8283. In addition, a dummy variable was created to indicate whether or not the taxpayer paid estimated taxes.

Self Preparers with Software

$$\ln(time)_k = b_0 + b_1 OT_k + b_2 LSM_k + b_3 MFJ_k + b_4 MFS_k + b_5 AI_k + b_6 TF_k + b_7 NS_k + b_8 MPM_k + e_k$$

$$(k = 1,2,...,K)$$

OT = indicator for Owing Taxes

LSM = indicator for Legal Status Minor

MFJ = indicator for Married Filing Jointly

MFS = indicator for Married Filing Separately

AI = Attribute Index

TF = indicator for TeleFile

NS = indicator for Spent Time on Forms Not Submitted

MPM = indicator for Multiple Preparation Methods

The final specifications and results for each econometric equation are presented in Tables C.1 through C.6 (Appendix C).

IV.3.2.5 REGRESSION DIAGNOSTICS - MULTICOLLINEARITY

When an independent variable is nearly a linear combination of other independent variables in a regression model, the affected parameter estimates are unstable and have high standard errors, a problem referred to as *collinearity* or *multicollinearity*. Multicollinearity was a particular concern in our estimation, as many of the factors that influence taxpayer burden are highly correlated with one another. Most notably, the 20 attribute scores discussed in Section IV.2.3.3 will generally increase or decrease in tandem, depending on the number of forms and schedules a taxpayer files.

There are several potential remedies for multicollinearity, including: (1) imposing restrictions on the parameters based on non-sample information, (2) employing factor analysis, and (3) employing principal components. In our final estimation of the econometric equations, we relied on principal components analysis to alleviate the problem of multicollinearity. The purpose of principal component analysis is to derive a small number of variables (principal components) from a larger set of variables while still retaining as much of the information in the original larger set of variables as possible.[27] Often a small number of principal components can be used in place of the original variables. In our application,

[27] Greene, William(1993), *Econometric Analysis,* New York: Macmillan Publishing Co.

information from the first principal component was used to construct a unidimensional attribute index that was used in place of the 20 attribute scores.

We chose not to impose a priori restrictions on the parameters with non-sample information, since the theoretical and empirical literature offered little guidance regarding what the restrictions might be. Additionally, we chose not to employ factor analysis, since the main purpose of factor analysis is to explain the correlations or covariances among a set of variables in terms of a limited number of unobservable, latent variables.[28] These latent variables are not generally computable as linear combinations of the original variables. The latent variable structure itself, is often guided by a priori theories, or by the researcher, and it describes how the larger set of variables interact with one another. In our application, we had no reason to suspect that the larger set of variables relate to a latent set of variables.

The approach we adopted for diagnosing multicollinearity follows that of Belsley, Kuh, and Welsch (1980).[29] Belsey, Kuh, and Welsch (1980) suggest that, when the condition index approaches 10, weak dependencies may be starting to affect the regression estimates.[30] When this index is larger than 100, the estimated coefficients and standard errors may have a fair amount of numerical error (although the statistical standard error almost always is much greater than the numerical error).[31] After replacing raw attribute scores with the attribute index, none of the econometric equations had a condition index in excess of 8.2.

IV.3.2.6 REGRESSION DIAGNOSTICS - HETEROSCEDASTICITY

A typical regression model is specified as:

$$y_i = bx_i + e_i$$

where the error term, e_i, is assumed to be identically and independently distributed, with a mean of zero, and a constant variance. If the error terms are not independent, or their variances are not constant, the estimated parameters remain unbiased, but the estimate of the covariance matrix is inconsistent. Various tests can help identify the presence of heteroscedasticity. The test we chose was developed by White

[28] Mulaik, S.A. (1972), *The Foundations of Factor Analysis*, New York: McGraw-Hill Book Co.

[29] Belsley, D.A., Kuh, E., and Welsch, R.E. (1980), *Regression Diagnostics*, New York: John Wiley & Sons, Inc.

[30] The condition index is the square root of the ratio of the largest eigenvalue to the smallest eigenvalue.

[31] The diagnostic results for multicollinearity can be found in *Phase B: Wage and Investment Taxpayer Burden Study Group, Final Estimation Summary*, June 2001.

(1980).[32] Other tests were considered, including Breusch-Pagan, however, since or analysis did not uncover which variable(s) may be the cause of the heteroscedastic ity, we opted for the more general test devised by White (1980), which makes no assumption about the functional form of the heteroscedasticity, nor the variables that are causing it.[33]

We found some degree of heteroscedasticity to be present in nearly all of our initial econometric equations. After reviewing the test results, we corrected the standard errors using the methodology proposed by White (1980). In general, the variables that were significant using the uncorrected standard errors, remained significant after the correction. Thus, we determined that heteroscedasticity did not have a large impact on the specification of the equations or the importance of key variables in the equations.

IV.3.2.7 LEVEL OF BURDEN – SEGMENTATION MODELS

Using the criteria described in Section IV.3.2.1, we identified several lower-priority burden categories, including: (1) form submission time, and (2) all types of out-of-pocket burden except for money paid to a professional for return preparation. These lower-priority types of burden were estimated using segmentation models.

The methodology used to estimate these segmentation models is identical to that described in Section IV.3.2.2 (Burden Incidence), except that the outcome variable here is the mean level of burden rather than the probability of experiencing burden of a given type.[34]

[32] White, H. (1980), "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test For Heteroskedasticity," *Econometrics*, 48, 817-838.

[33] The detailed heteroscedasticity tests and the corrected standard errors can be found in *Phase B: Wage and Investment Taxpayer Burden Study Group, Final Estimation Summary*, June 2001.

[34] The detailed segmentation models can be found in *Phase B: Wage and Investment Taxpayer Burden Study Group, Final Estimation Summary*, June 2001.

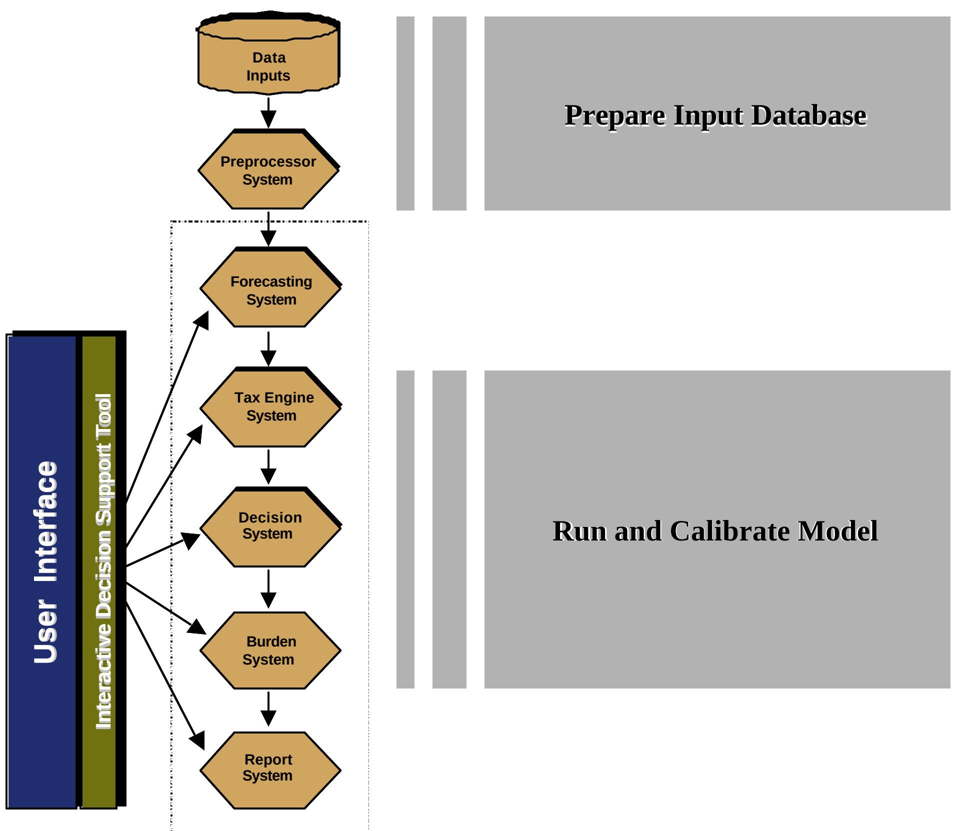
V. MODEL DEVELOPMENT

The model development phase involves three distinct tasks: (1) develop a production data file to drive the model, (2) establish a set of simulation levers that allow users to change parameters or alter behavioral assumptions, and (3) design and implement a software tool that accommodates user- and data-inputs and applies a series of algorithms to simulate outcomes. In this section, we discuss the first two model development tasks.[35]

V.1 DEVELOPING A PRODUCTION DATA FILE

Developing a production data file entails several steps, which roughly follow the modeling sequence discussed in Section II.3. As illustrated in ***Figure V-1***, these steps can be broadly categorized into two groups: (1) prepare an input database and (2) run the model and calibrate the predicted outcomes.

FIGURE V-1: PRODUCTION FILE DEVELOPMENT PROCESS



[35] For more information on software design, see the Wage and Investment Burden Model Programmer's Guide. For a detailed overview of the graphical user interface, including screen shots, see the Wage and Investment Burden Model User's Guide.

V.1.1 PREPARE AN INPUT DATABASE

The primary source of input data for the burden model is the Continuous Work History Sample (CWHS)—a simple random sample of tax returns prepared annually by IRS's Statistics of Income Division. The CWHS data file provided by IRS contains 63,435 records, of which 46,120 represent TY99 W&I returns (the population of interest). The data file contains information on 872 variables, including 23 variables that represent prior year (i.e., TY98) filing outcomes.[36] Many of these variables are not needed for the W&I burden model (e.g., those that pertain to Self-Employed tax forms and schedules) and therefore are deleted as part of the file development process.

After excluding unnecessary observations and variables, the resulting CWHS data file is passed through a Preprocessor module to impute or infer a number of data elements and to reformat certain variables for use in the model (e.g., converting year-of-birth to age). Finally, selected data elements from the taxpayer survey are merged to the CWHS data file through a "constrained match" process.[37] The methodology for this constrained match is similar to hot-deck imputation—donor records from the taxpayer survey donate their characteristics to host records on the CWHS file based on proximity of auxiliary characteristics.[38] The constrained match approach offers the additional benefit of preserving the distribution of the host data file across user-defined strata, which are used to constrain the set of donor records that can be chosen for a given host record.

V.1.2 CALIBRATE THE MODEL

Once the input data file is prepared, it is passed through each component of the burden module to generate simulated outcomes (e.g., filing outcomes, compliance methods, burden levels).[39] The simulated outcomes for each taxpayer are then compared against reported outcomes, and the model is calibrated to minimize the impact of any discrepancies. The nature of this calibration depends on the reason for the discrepancy and on our ability to resolve the discrepancy, as described below.

[36] For detailed information on the CWHS sample, see the report, Data Dictionary and File Layout – FINAL (October 31, 2001).

[37] Among the variables merged from the survey data are, (1) reported burden levels, and (2) burden determinants that are not present on the CWHS. Reported burden levels are used to calibrate the model-simulated burden levels, as described in Section V.1.3.

[38] It should be noted that some of the auxiliary characteristics used to identify donor records are filing outcomes. Consequently, it is necessary to run the Tax Engine prior to the constrained match, so that the filing outcomes used as auxiliary characteristics are consistent with those simulated by the model.

[39] For this initial "calibration" run of the model, data are not aged, and all of the model parameters and code reflect conditions in the data year.

V.1.2.1 CALIBRATING THE TAX ENGINE

Most discrepancies in the Tax Engine (i.e., differences between simulated and reported filing outcomes) stem from one of three problems. The first is incomplete input data, a problem that arises when the CWHS data file lacks one ore more data elements that are needed to accurately simulate filing outcomes.[40] In many cases, we can deduce that the discrepancy is caused by a single missing data element, which allows us to correct the discrepancy by computing the missing element in the pre-processor. In other cases, we cannot reliably ascribe the discrepancy to any single source, and therefore cannot correct it.

The second source of discrepancies in the Tax Engine is internally inconsistent input data. Most of these inconsistencies involve a tax return whose total income is not equal to the sum of all taxable income items on that return. In general, we did not perform any calibration to resolve these data-related discrepancies, as the solution involves judgmental data cleaning. Due to the prevalence of discrepancies in total income, however, we chose to balance this field through the use of a residual income field, which forces total income to equal the sum of its components.

The final source of discrepancies is that the Tax Engine algorithms—which seek to minimize tax liability—can overstate participation rates in optional credits. In most cases, we resolved this problem by calibrating simulated participation to align with participation as reported in the data.

V.1.2.2 CALIBRATING THE DECISION SYSTEM

Calibration in the Decision System is unnecessary. The model is not designed to simulate decision outcomes in the base year (TY99 in this case) and therefore cannot simulate an outcome that is at odds with a taxpayer-reported outcome. Rather, the model predicts the probability of each possible decision outcome in the base year. These base-year probabilities are, however, used to modify probabilities under what-if scenarios. For example, if a taxpayer reports a base-year preparation method that the model views as highly unlikely, this information will be used to predict the taxpayer's preparation method under what-if scenarios—thereby enforcing the taxpayer's unexplained tendency towards the chosen preparation method.[41]

[40] For example, the simulated tax amount may be less than taxpayer-reported tax amount because the CWHS lacks information on one of the Other Taxes on Form 1040, Line 56.

[41] For a more detailed discussion of the statistical model underlying the Decision System, see Section IV.3.1.

V.1.2.3 Calibrating the Burden System

Discrepancies in the burden system (i.e., differences between reported and predicted burden levels) reflect the fact that our econometric and statistical models do not explain all of the variability in taxpayer burden. The model resolves these discrepancies through the use of residuals and distributional indices that force the predicted burden to match reported burden in the base year.[42] It is important to note that this calibration does not affect model-based estimates of change in burden—it merely ties the base-year prediction of burden to the level observed in the taxpayer survey, which is the best available benchmark.

V.2 Simulation Levers

Having reviewed the process of creating a production data file, we are now prepared to discuss use of the production model—and in particular, the simulation levers that are available to model users. The user interface provides access to six groups of simulation levers, which are listed below, and described in greater detail in Sections V.2.1 through V.2.6.[43]

- **Forecast Parameters**, which control the way input data and selected tax system parameters are aged
- **Tax Law Parameters** and **Tax Law Structure**, which represent the numeric constants and tax rules that determine tax liability
- **Burden Attributes**, which describe the activities associated with existing filing outcomes
- **Decision Parameters**, which allow the user to adjust model-predicted outcomes with regard to preparation method and submission method
- **Administrative Parameters**, which allow the user to adjust model-predicted outcomes with regard to the use of IRS services and the methods used to gather tax materials.

V.2.1 Forecast Parameters

In order to forecast taxpayer burden in a future year, it is first necessary to age the data to reflect future conditions. Even current year simulations generally require aging, due to the inevitable gap between the data year and the year the model is used. To age the data, the model uses 41 different forecast parameters,

[42] In the econometric models, predicted burden is adjusted by a residual that is equal to the discrepancy between reported burden and base-year predicted burden. In the statistical models, each observation is assigned a number between 0 and 1 that represents its location on an exponential distribution of burden for the corresponding taxpayer segment.

[43] Scenarios that cannot be implemented through the user interface can generally be implemented through a combination of changes to the model code or the input data. For more information on implementing these scenarios, see the Wage and Investment Burden Model Programmer's Guide.

which account for factors such as inflation, real income growth, and changes to the demographic composition of the population. Users can specify changes to any of the following types of forecast parameters.

- **CPI-U August 31.** This parameter is used for federal indexing of selected tax system parameters (e.g., exemptions, floors, ceilings, phase-out ranges). It is constructed as the urban average Consumer Price Index (CPI) for the twelve-month period from September through August.

- **CPI-U End of Year.** This parameter is used as a multiplier for all income and expense items to account for nominal changes in value due to inflation. It is equal to the urban average CPI for the twelve-month period from January through December.

- **Real per capita income and expense growth rates.** These parameters are used as multipliers for selected income and expense items to account for real per capita economic growth. Separate parameters (multipliers) are used for several distinct income types to allow for differential growth rates.

- **Return growth rates by AGI class:** These parameters are used to account for year-to-year changes in the distribution of tax returns due to population growth and new entrants into the tax system. Separate parameters are used for each of the 15 defined AGI classes.

The complete forecasting system works as follows. First, weight factors are computed to reflect changes in the distribution of taxpayers across AGI categories, resulting from user-specified rates for return growth and real per capita income growth. Second, income and expense amounts in each base-year tax return are scaled up to reflect both nominal and real growth rates.

V.2.2 TAX LAW PARAMETERS

Tax Law Parameters are the values that, when combined with Tax Law Structures, establish the rules used to determine tax liability. To help illustrate the difference between Tax Law Parameters and Tax Law Structure, consider the deduction for contributions to an Individual Retirement Arrangement (IRA). The Tax Law Structure surrounding the IRA deduction specifies that a taxpayer may claim a deduction against his total income for contributions made to selected savings accounts, provided that he qualifies based on his income, filing status, and coverage by work-based retirement plans. Tax Law Parameters are the numeric constants associated with this provision. For example, the maximum deduction for an IRA contribution in TY99 is $2,000, and single taxpayers who are covered by a work-based retirement are eligible for a reduced deduction if their modified AGI exceeds $31,000. The purpose of Tax Law Parameters is to give users a simple tool for modifying existing tax laws. Specifically, users have access to the following simulation levers:

- **Policy Parameters.** These parameters are numeric elements of the statutory rules that determine tax liability. Examples of policy parameters include the two IRA parameters listed above, the marginal tax rates for each income bracket, and the 7.5 percent (of AGI) exclusion for deduction of medical and dental expenses.

- **Administrative Parameters.** These parameters are numeric elements of administrative rules, often relating to the documentation of income or expenses. Examples include the $400 threshold for documentation of interest and dividend income, and the $500 threshold for documentation of charitable contributions.

- **Indexing and Phase-In Rules.** Some tax system parameters are indexed or phased-in, so that their values increase according to the inflation rate or some other statutory schedule. To fully specify the level of an indexed parameter, two values are needed—the base year for the indexed parameter and the value of the parameter in the base year. The model uses these values in conjunction with the CPI-U August 31 parameter (see Section V.2.1) to determine the value of the corresponding tax system parameter in the simulation year. Examples of indexed tax system parameters include the standard deduction (1987 base year), the personal exemption (1988 base year), and income thresholds for EITC eligibility (1995 base year).

V.2.3 TAX LAW STRUCTURES

The purpose of the Tax Law Structure lever is to allow users to simulate the effect of new structures, or rules, in the tax system. Unlike tax law parameters, which are easily specified as numeric constants, tax law structures are complex, multi-dimensional, and non-uniform. To accommodate all of the characteristics of a new tax law structure, the user interface would have to be extremely complicated and burdensome to use.

To avoid this complexity, the working group decided that the model should accommodate new tax structures through the use of supplemental data inputs. Under this approach, a user simply creates a new data field that indicates (for each taxpayer) the amount associated with the new tax structure, then imports that supplemental data field into the model for further processing. Users can modify four categories of tax structures—income elements, adjustments to income, itemized deductions, and credits. In each category, the user can:

- **Deactivate an existing tax structure.** The user can specify which of the existing tax structures are active in a given scenario. For example, this feature could be used to simulate the impact of eliminating an itemized deduction or credit.

- **Create a new tax structure.** The user can create a new structure by reading in a supplemental data field that specifies (for each taxpayer) the amount associated with the new structure.

- **Assign burden attributes to a new structure.** An essential step in creating a new tax structure is to specify the burden attributes of the new structure. The model uses these attributes, which describe the activities associated with a tax structure, to predict the structure's impact on taxpayer burden.

V.2.4 BURDEN ATTRIBUTES

As discussed in Section IV.2.3.3, burden attributes are characteristics of filing outcomes. The attribute profile of a filing outcome enumerates the activities associated with that outcome and attempts to quantify the complexity of those activities. To simulate the impact of changes to existing tax forms, model users can adjust the attributes of those forms. For example, if IRS streamlined a worksheet to eliminate several calculations, a model user could reduce the number of "Calculate" attributes associated with specific line items on that worksheet, then run the model to determine the impact on taxpayer burden.

V.2.5 DECISION PARAMETERS

While most taxpayers exhibit inertial behavior with respect to preparation method (i.e., it does not change from year to year), there are several factors that are empirically correlated with changes in preparation method. The model enforces these relationships through the use of predictive equations that simulate each taxpayer's choice of preparation method, as well as the related choice of submission method. However, both of these choices are likely to be influenced by strong exogenous factors in future years. One obvious example of such a factor is the IRS initiative to increase access to free electronic filing services. Users can specify the impact of these exogenous changes through the use of Decision Parameters.

Specifically, users can adjust the model-predicted preparation and submission methods through the use of transition matrices. In effect, these transition matrices take all of the taxpayers predicted to use a given preparation method and redistribute them to match a user-specified probability distribution. By default, each transition matrix has values of 100 percent along the diagonal and 0 percent in all other cells—resulting in no adjustment to the model-predicted outcomes. By adjusting the percentages in each cell, users can impose anything from a marginal shift away from the predicted distribution, to a complete override with user-specified outcomes.

V.2.6 ADMINISTRATIVE PARAMETERS

In meeting their filing requirements, taxpayers face a number of choices. In addition to deciding how to prepare and submit their returns, taxpayers must decide how they will obtain tax materials and whether they will use IRS-provided taxpayer services. The latter two decisions are particularly important from IRS's perspective, as they represent the most direct interaction the Service has with most W&I

taxpayers—and an opportunity to reduce taxpayer burden through effective administration of the tax system.

Administrative parameters allow model users to adjust the model's predictions regarding the use of IRS services and the methods that taxpayers use to gather tax materials (GTM). Specifically, users can change:

- **Use of IRS Services.** Users can specify the percentage of taxpayers that use each of the following IRS services: (1) toll-free tax assistance; (2) VITA, TCE, or IRS walk-in site; (3) IRS web site; and (4) TeleTax. Separate percentages can be specified for taxpayers using each of the three possible preparation methods.

- **Wait Time for IRS Services.** Users can specify two separate parameters for each IRS service: (1) the percentage of service users that experienced a non-zero wait time, and (2) the average wait time among taxpayers that experienced a non-zero wait time.

- **Incidence of GTM Methods.** Users can specify the percentage of taxpayers that gathered tax materials using each of the following methods: (1) picked-up materials from a library or post office, (2) downloaded or ordered materials from the IRS, or (3) purchased tax books or guides.

APPENDIX A: DECISION EQUATIONS

Table A.1
Decision Equation Specifications: Preparation Method

Preparation Method in 1998	Change in Primary Form	Net Change in Non-Primary Form	Preparation Method in 1999	Observed Probability	Predicted Probability
Paid-Prep	1040 to 1040A	No Change	Paid-Prep	0.823	0.874
Paid-Prep	1040 to 1040A	No Change	Self-Prep w/o SW	0.097	0.074
Paid-Prep	1040 to 1040A	No Change	Self-Prep w/SW	0.080	0.052
Paid-Prep	1040 to 1040A	Increase	Paid-Prep	0.868	0.927
Paid-Prep	1040 to 1040A	Increase	Self-Prep w/o SW	0.076	0.025
Paid-Prep	1040 to 1040A	Increase	Self-Prep w/SW	0.057	0.048
Paid-Prep	1040 to 1040A	Decrease	Paid-Prep	0.724	0.868
Paid-Prep	1040 to 1040A	Decrease	Self-Prep w/o SW	0.181	0.075
Paid-Prep	1040 to 1040A	Decrease	Self-Prep w/SW	0.095	0.058
Paid-Prep	1040 to 1040EZ	No Change	Paid-Prep	0.405	0.473
Paid-Prep	1040 to 1040EZ	No Change	Self-Prep w/o SW	0.524	0.492
Paid-Prep	1040 to 1040EZ	No Change	Self-Prep w/SW	0.071	0.036
Paid-Prep	1040 to 1040EZ	Increase	Paid-Prep	0.800	0.718
Paid-Prep	1040 to 1040EZ	Increase	Self-Prep w/o SW	0.200	0.235
Paid-Prep	1040 to 1040EZ	Increase	Self-Prep w/SW	0.000	0.048
Paid-Prep	1040 to 1040EZ	Decrease	Paid-Prep	0.273	0.468
Paid-Prep	1040 to 1040EZ	Decrease	Self-Prep w/o SW	0.727	0.493
Paid-Prep	1040 to 1040EZ	Decrease	Self-Prep w/SW	0.000	0.040
Paid-Prep	1040A to 1040	No Change	Paid-Prep	0.878	0.950
Paid-Prep	1040A to 1040	No Change	Self-Prep w/o SW	0.039	0.018
Paid-Prep	1040A to 1040	No Change	Self-Prep w/SW	0.082	0.033
Paid-Prep	1040A to 1040	Increase	Paid-Prep	0.857	0.965
Paid-Prep	1040A to 1040	Increase	Self-Prep w/o SW	0.045	0.006
Paid-Prep	1040A to 1040	Increase	Self-Prep w/SW	0.098	0.029
Paid-Prep	1040A to 1040	Decrease	Paid-Prep	0.829	0.946
Paid-Prep	1040A to 1040	Decrease	Self-Prep w/o SW	0.073	0.018
Paid-Prep	1040A to 1040	Decrease	Self-Prep w/SW	0.098	0.036
Paid-Prep	1040A to 1040EZ	No Change	Paid-Prep	0.469	0.663
Paid-Prep	1040A to 1040EZ	No Change	Self-Prep w/o SW	0.510	0.313
Paid-Prep	1040A to 1040EZ	No Change	Self-Prep w/SW	0.020	0.024
Paid-Prep	1040A to 1040EZ	Increase	Paid-Prep	1.000	0.847
Paid-Prep	1040A to 1040EZ	Increase	Self-Prep w/o SW	0.000	0.126
Paid-Prep	1040A to 1040EZ	Increase	Self-Prep w/SW	0.000	0.027
Paid-Prep	1040A to 1040EZ	Decrease	Paid-Prep	0.542	0.658
Paid-Prep	1040A to 1040EZ	Decrease	Self-Prep w/o SW	0.458	0.315
Paid-Prep	1040A to 1040EZ	Decrease	Self-Prep w/SW	0.000	0.027
Paid-Prep	1040EZ to 1040	No Change	Paid-Prep	0.842	0.940
Paid-Prep	1040EZ to 1040	No Change	Self-Prep w/o SW	0.064	0.010
Paid-Prep	1040EZ to 1040	No Change	Self-Prep w/SW	0.094	0.050
Paid-Prep	1040EZ to 1040	Increase	Paid-Prep	0.864	0.952
Paid-Prep	1040EZ to 1040	Increase	Self-Prep w/o SW	0.042	0.003
Paid-Prep	1040EZ to 1040	Increase	Self-Prep w/SW	0.093	0.045
Paid-Prep	1040EZ to 1040	Decrease	Paid-Prep	1.000	0.934
Paid-Prep	1040EZ to 1040	Decrease	Self-Prep w/o SW	0.000	0.010
Paid-Prep	1040EZ to 1040	Decrease	Self-Prep w/SW	0.000	0.056

Table A.1
Decision Equation Specifications: Preparation Method

Preparation Method in 1998	Change in Primary Form	Net Change in Non-Primary Form	Preparation Method in 1999	Observed Probability	Predicted Probability
Paid-Prep	1040EZ to 1040A	No Change	Paid-Prep	0.777	0.926
Paid-Prep	1040EZ to 1040A	No Change	Self-Prep w/o SW	0.138	0.035
Paid-Prep	1040EZ to 1040A	No Change	Self-Prep w/SW	0.085	0.039
Paid-Prep	1040EZ to 1040A	Increase	Paid-Prep	0.839	0.953
Paid-Prep	1040EZ to 1040A	Increase	Self-Prep w/o SW	0.075	0.011
Paid-Prep	1040EZ to 1040A	Increase	Self-Prep w/SW	0.086	0.035
Paid-Prep	1040EZ to 1040A	Decrease	Paid-Prep	0.667	0.921
Paid-Prep	1040EZ to 1040A	Decrease	Self-Prep w/o SW	0.333	0.036
Paid-Prep	1040EZ to 1040A	Decrease	Self-Prep w/SW	0.000	0.044
Paid-Prep	No Change	No Change	Paid-Prep	0.907	0.887
Paid-Prep	No Change	No Change	Self-Prep w/o SW	0.051	0.060
Paid-Prep	No Change	No Change	Self-Prep w/SW	0.043	0.053
Paid-Prep	No Change	Increase	Paid-Prep	0.937	0.931
Paid-Prep	No Change	Increase	Self-Prep w/o SW	0.014	0.020
Paid-Prep	No Change	Increase	Self-Prep w/SW	0.049	0.049
Paid-Prep	No Change	Decrease	Paid-Prep	0.890	0.881
Paid-Prep	No Change	Decrease	Self-Prep w/o SW	0.050	0.061
Paid-Prep	No Change	Decrease	Self-Prep w/SW	0.060	0.059
Self-Prep w/o SW	1040 to 1040A	No Change	Paid-Prep	0.091	0.069
Self-Prep w/o SW	1040 to 1040A	No Change	Self-Prep w/o SW	0.800	0.851
Self-Prep w/o SW	1040 to 1040A	No Change	Self-Prep w/SW	0.109	0.080
Self-Prep w/o SW	1040 to 1040A	Increase	Paid-Prep	0.258	0.171
Self-Prep w/o SW	1040 to 1040A	Increase	Self-Prep w/o SW	0.621	0.657
Self-Prep w/o SW	1040 to 1040A	Increase	Self-Prep w/SW	0.121	0.172
Self-Prep w/o SW	1040 to 1040A	Decrease	Paid-Prep	0.108	0.068
Self-Prep w/o SW	1040 to 1040A	Decrease	Self-Prep w/o SW	0.703	0.844
Self-Prep w/o SW	1040 to 1040A	Decrease	Self-Prep w/SW	0.189	0.088
Self-Prep w/o SW	1040 to 1040EZ	No Change	Paid-Prep	0.068	0.007
Self-Prep w/o SW	1040 to 1040EZ	No Change	Self-Prep w/o SW	0.824	0.984
Self-Prep w/o SW	1040 to 1040EZ	No Change	Self-Prep w/SW	0.108	0.010
Self-Prep w/o SW	1040 to 1040EZ	Increase	Paid-Prep	0.000	0.020
Self-Prep w/o SW	1040 to 1040EZ	Increase	Self-Prep w/o SW	1.000	0.954
Self-Prep w/o SW	1040 to 1040EZ	Increase	Self-Prep w/SW	0.000	0.026
Self-Prep w/o SW	1040 to 1040EZ	Decrease	Paid-Prep	0.140	0.006
Self-Prep w/o SW	1040 to 1040EZ	Decrease	Self-Prep w/o SW	0.721	0.983
Self-Prep w/o SW	1040 to 1040EZ	Decrease	Self-Prep w/SW	0.140	0.011
Self-Prep w/o SW	1040A to 1040	No Change	Paid-Prep	0.373	0.229
Self-Prep w/o SW	1040A to 1040	No Change	Self-Prep w/o SW	0.437	0.619
Self-Prep w/o SW	1040A to 1040	No Change	Self-Prep w/SW	0.190	0.152
Self-Prep w/o SW	1040A to 1040	Increase	Paid-Prep	0.385	0.411
Self-Prep w/o SW	1040A to 1040	Increase	Self-Prep w/o SW	0.410	0.350
Self-Prep w/o SW	1040A to 1040	Increase	Self-Prep w/SW	0.205	0.239
Self-Prep w/o SW	1040A to 1040	Decrease	Paid-Prep	0.360	0.223
Self-Prep w/o SW	1040A to 1040	Decrease	Self-Prep w/o SW	0.440	0.611
Self-Prep w/o SW	1040A to 1040	Decrease	Self-Prep w/SW	0.200	0.166
Self-Prep w/o SW	1040A to 1040EZ	No Change	Paid-Prep	0.134	0.014
Self-Prep w/o SW	1040A to 1040EZ	No Change	Self-Prep w/o SW	0.805	0.976
Self-Prep w/o SW	1040A to 1040EZ	No Change	Self-Prep w/SW	0.061	0.010

Table A.1
Decision Equation Specifications: Preparation Method

Preparation Method in 1998	Change in Primary Form	Net Change in Non-Primary Form	Preparation Method in 1999	Observed Probability	Predicted Probability
Self-Prep w/o SW	1040A to 1040EZ	Increase	Paid-Prep	0.000	0.043
Self-Prep w/o SW	1040A to 1040EZ	Increase	Self-Prep w/o SW	1.000	0.930
Self-Prep w/o SW	1040A to 1040EZ	Increase	Self-Prep w/SW	0.000	0.027
Self-Prep w/o SW	1040A to 1040EZ	Decrease	Paid-Prep	0.083	0.014
Self-Prep w/o SW	1040A to 1040EZ	Decrease	Self-Prep w/o SW	0.917	0.975
Self-Prep w/o SW	1040A to 1040EZ	Decrease	Self-Prep w/SW	0.000	0.011
Self-Prep w/o SW	1040EZ to 1040	No Change	Paid-Prep	0.453	0.282
Self-Prep w/o SW	1040EZ to 1040	No Change	Self-Prep w/o SW	0.291	0.428
Self-Prep w/o SW	1040EZ to 1040	No Change	Self-Prep w/SW	0.256	0.290
Self-Prep w/o SW	1040EZ to 1040	Increase	Paid-Prep	0.457	0.420
Self-Prep w/o SW	1040EZ to 1040	Increase	Self-Prep w/o SW	0.164	0.200
Self-Prep w/o SW	1040EZ to 1040	Increase	Self-Prep w/SW	0.379	0.380
Self-Prep w/o SW	1040EZ to 1040	Decrease	Paid-Prep	0.222	0.271
Self-Prep w/o SW	1040EZ to 1040	Decrease	Self-Prep w/o SW	0.667	0.417
Self-Prep w/o SW	1040EZ to 1040	Decrease	Self-Prep w/SW	0.111	0.313
Self-Prep w/o SW	1040EZ to 1040A	No Change	Paid-Prep	0.305	0.136
Self-Prep w/o SW	1040EZ to 1040A	No Change	Self-Prep w/o SW	0.621	0.753
Self-Prep w/o SW	1040EZ to 1040A	No Change	Self-Prep w/SW	0.074	0.111
Self-Prep w/o SW	1040EZ to 1040A	Increase	Paid-Prep	0.392	0.290
Self-Prep w/o SW	1040EZ to 1040A	Increase	Self-Prep w/o SW	0.431	0.503
Self-Prep w/o SW	1040EZ to 1040A	Increase	Self-Prep w/SW	0.177	0.207
Self-Prep w/o SW	1040EZ to 1040A	Decrease	Paid-Prep	0.286	0.133
Self-Prep w/o SW	1040EZ to 1040A	Decrease	Self-Prep w/o SW	0.714	0.745
Self-Prep w/o SW	1040EZ to 1040A	Decrease	Self-Prep w/SW	0.000	0.122
Self-Prep w/o SW	No Change	No Change	Paid-Prep	0.068	0.084
Self-Prep w/o SW	No Change	No Change	Self-Prep w/o SW	0.841	0.820
Self-Prep w/o SW	No Change	No Change	Self-Prep w/SW	0.091	0.096
Self-Prep w/o SW	No Change	Increase	Paid-Prep	0.203	0.196
Self-Prep w/o SW	No Change	Increase	Self-Prep w/o SW	0.598	0.606
Self-Prep w/o SW	No Change	Increase	Self-Prep w/SW	0.199	0.199
Self-Prep w/o SW	No Change	Decrease	Paid-Prep	0.070	0.082
Self-Prep w/o SW	No Change	Decrease	Self-Prep w/o SW	0.824	0.813
Self-Prep w/o SW	No Change	Decrease	Self-Prep w/SW	0.106	0.106
Self-Prep w/SW	1040 to 1040A	No Change	Paid-Prep	0.232	0.143
Self-Prep w/SW	1040 to 1040A	No Change	Self-Prep w/o SW	0.188	0.119
Self-Prep w/SW	1040 to 1040A	No Change	Self-Prep w/SW	0.580	0.739
Self-Prep w/SW	1040 to 1040A	Increase	Paid-Prep	0.172	0.172
Self-Prep w/SW	1040 to 1040A	Increase	Self-Prep w/o SW	0.207	0.045
Self-Prep w/SW	1040 to 1040A	Increase	Self-Prep w/SW	0.621	0.783
Self-Prep w/SW	1040 to 1040A	Decrease	Paid-Prep	0.140	0.130
Self-Prep w/SW	1040 to 1040A	Decrease	Self-Prep w/o SW	0.190	0.110
Self-Prep w/SW	1040 to 1040A	Decrease	Self-Prep w/SW	0.670	0.759
Self-Prep w/SW	1040 to 1040EZ	No Change	Paid-Prep	0.071	0.056
Self-Prep w/SW	1040 to 1040EZ	No Change	Self-Prep w/o SW	0.684	0.574
Self-Prep w/SW	1040 to 1040EZ	No Change	Self-Prep w/SW	0.245	0.370
Self-Prep w/SW	1040 to 1040EZ	Increase	Paid-Prep	0.000	0.100
Self-Prep w/SW	1040 to 1040EZ	Increase	Self-Prep w/o SW	1.000	0.321
Self-Prep w/SW	1040 to 1040EZ	Increase	Self-Prep w/SW	0.000	0.579

Table A.1
Decision Equation Specifications: Preparation Method

Preparation Method in 1998	Change in Primary Form	Net Change in Non-Primary Form	Preparation Method in 1999	Observed Probability	Predicted Probability
Self-Prep w/SW	1040 to 1040EZ	Decrease	Paid-Prep	0.023	0.053
Self-Prep w/SW	1040 to 1040EZ	Decrease	Self-Prep w/o SW	0.791	0.552
Self-Prep w/SW	1040 to 1040EZ	Decrease	Self-Prep w/SW	0.186	0.395
Self-Prep w/SW	1040A to 1040	No Change	Paid-Prep	0.288	0.240
Self-Prep w/SW	1040A to 1040	No Change	Self-Prep w/o SW	0.102	0.044
Self-Prep w/SW	1040A to 1040	No Change	Self-Prep w/SW	0.610	0.716
Self-Prep w/SW	1040A to 1040	Increase	Paid-Prep	0.338	0.272
Self-Prep w/SW	1040A to 1040	Increase	Self-Prep w/o SW	0.054	0.016
Self-Prep w/SW	1040A to 1040	Increase	Self-Prep w/SW	0.608	0.713
Self-Prep w/SW	1040A to 1040	Decrease	Paid-Prep	0.467	0.220
Self-Prep w/SW	1040A to 1040	Decrease	Self-Prep w/o SW	0.133	0.041
Self-Prep w/SW	1040A to 1040	Decrease	Self-Prep w/SW	0.400	0.739
Self-Prep w/SW	1040A to 1040EZ	No Change	Paid-Prep	0.200	0.113
Self-Prep w/SW	1040A to 1040EZ	No Change	Self-Prep w/o SW	0.500	0.526
Self-Prep w/SW	1040A to 1040EZ	No Change	Self-Prep w/SW	0.300	0.361
Self-Prep w/SW	1040A to 1040EZ	Decrease	Paid-Prep	0.067	0.107
Self-Prep w/SW	1040A to 1040EZ	Decrease	Self-Prep w/o SW	0.667	0.507
Self-Prep w/SW	1040A to 1040EZ	Decrease	Self-Prep w/SW	0.267	0.386
Self-Prep w/SW	1040EZ to 1040	No Change	Paid-Prep	0.292	0.174
Self-Prep w/SW	1040EZ to 1040	No Change	Self-Prep w/o SW	0.083	0.018
Self-Prep w/SW	1040EZ to 1040	No Change	Self-Prep w/SW	0.625	0.808
Self-Prep w/SW	1040EZ to 1040	Increase	Paid-Prep	0.462	0.196
Self-Prep w/SW	1040EZ to 1040	Increase	Self-Prep w/o SW	0.000	0.006
Self-Prep w/SW	1040EZ to 1040	Increase	Self-Prep w/SW	0.539	0.798
Self-Prep w/SW	1040EZ to 1040	Decrease	Paid-Prep	0.000	0.158
Self-Prep w/SW	1040EZ to 1040	Decrease	Self-Prep w/o SW	1.000	0.017
Self-Prep w/SW	1040EZ to 1040	Decrease	Self-Prep w/SW	0.000	0.825
Self-Prep w/SW	1040EZ to 1040A	No Change	Paid-Prep	0.200	0.198
Self-Prep w/SW	1040EZ to 1040A	No Change	Self-Prep w/o SW	0.000	0.074
Self-Prep w/SW	1040EZ to 1040A	No Change	Self-Prep w/SW	0.800	0.728
Self-Prep w/SW	1040EZ to 1040A	Increase	Paid-Prep	0.500	0.230
Self-Prep w/SW	1040EZ to 1040A	Increase	Self-Prep w/o SW	0.000	0.027
Self-Prep w/SW	1040EZ to 1040A	Increase	Self-Prep w/SW	0.500	0.743
Self-Prep w/SW	1040EZ to 1040A	Decrease	Paid-Prep	0.000	0.181
Self-Prep w/SW	1040EZ to 1040A	Decrease	Self-Prep w/o SW	1.000	0.069
Self-Prep w/SW	1040EZ to 1040A	Decrease	Self-Prep w/SW	0.000	0.750
Self-Prep w/SW	No Change	No Change	Paid-Prep	0.131	0.146
Self-Prep w/SW	No Change	No Change	Self-Prep w/o SW	0.081	0.097
Self-Prep w/SW	No Change	No Change	Self-Prep w/SW	0.788	0.757
Self-Prep w/SW	No Change	Increase	Paid-Prep	0.159	0.173
Self-Prep w/SW	No Change	Increase	Self-Prep w/o SW	0.040	0.036
Self-Prep w/SW	No Change	Increase	Self-Prep w/SW	0.801	0.791
Self-Prep w/SW	No Change	Decrease	Paid-Prep	0.149	0.133
Self-Prep w/SW	No Change	Decrease	Self-Prep w/o SW	0.069	0.090
Self-Prep w/SW	No Change	Decrease	Self-Prep w/SW	0.782	0.777

Table A.2
Decision Equation Specifications: Submission Method
(for Prep Method = Self-Prep w/o Software)

Submission Method in 1998	Refund in 1999	Submission Method in 1999	Observed Probability	Predicted Probability
Paper	No	Paper	0.894	0.877
Paper	No	Telefile	0.106	0.123
Paper	Yes	Paper	0.753	0.755
Paper	Yes	Telefile	0.248	0.245
Telefile	No	Paper	0.235	0.308
Telefile	No	Telefile	0.765	0.692
Telefile	Yes	Paper	0.168	0.161
Telefile	Yes	Telefile	0.832	0.839

Note: Data are for those eligible for Telefile, this includes Self-Prep w/o SW, 1040EZ filers, all other Self-Prep w/o SW filers are categorized with paper submission.

Table A.3
Decision Equation Specifications: Submission Method
(for Prep Method = Paid Prep)

Submission Method in 1998	Refund in 1999	Submission Method in 1999	Observed Probability	Predicted Probability
Paper	No	Paper	0.951	0.958
Paper	No	Electronic	0.049	0.042
Paper	Yes	Paper	0.800	0.796
Paper	Yes	Electronic	0.200	0.204
Electronic	No	Paper	0.438	0.380
Electronic	No	Electronic	0.562	0.620
Electronic	Yes	Paper	0.088	0.094
Electronic	Yes	Electronic	0.912	0.906

Table A.4
Decision Equation Specifications: Submission Method
(for Prep Method = Self Prep w/Software)

Submission Method in 1998	Refund in 1999	Submission Method in 1999	Observed Probability	Predicted Probability
Paper	No	Paper	0.902	0.905
Paper	No	Electronic	0.098	0.095
Paper	Yes	Paper	0.661	0.660
Paper	Yes	Electronic	0.339	0.340
Electronic	No	Paper	0.442	0.420
Electronic	No	Electronic	0.558	0.580
Electronic	Yes	Paper	0.126	0.129
Electronic	Yes	Electronic	0.874	0.871

APPENDIX B: BURDEN ESTIMATION FRAMEWORK

Table B.1
Burden Estimation Framework - Time Burden

	Paid-Prep		Self-Prep w/o SW		Self-Prep w/SW	
	Outliers Included	**Outliers Removed**	**Outliers Included**	**Outliers Removed**	**Outliers Included**	**Outliers Removed**
Recordkeeping						
Type of Approach (R^2)	***Econometric (.082)***		***Econometric (.093)***		***Econometric (.154)***	
% Impacted	75.7%	75.6%	80.7%	80.5%	92.6%	92.5%
Mean Burden	8.21	6.96	5.97	5.07	8.24	7.59
Median Burden	2.00	2.00	2.00	2.00	3.00	3.00
Standard Deviation	21.00	12.48	12.77	8.94	15.50	12.37
Standard Error	0.69	0.41	0.42	0.27	0.57	0.41
Importance	**High**		**High**		**High**	
Gathering Tax Materials						
Type of Approach (R^2)	***Econometric (.026)***		***Econometric (.155)***		***Econometric (.020)***	
% Impacted	45.6%	45.3%	67.7%	67.6%	85.2%	85.1%
Mean Burden	2.46	1.86	1.86	1.47	1.91	1.74
Median Burden	1.00	1.00	1.00	1.00	1.00	1.00
Standard Deviation	6.30	2.90	5.07	2.09	2.83	1.96
Standard Error	0.27	0.12	0.23	0.09	0.14	0.12
Importance	**Medium**		**Medium**		**Medium**	
IRS Services						
Type of Approach (R^2)	***Econometric (.595)***		***Econometric (.550)***		***Econometric (.531)***	
% Impacted	10.3%	10.3%	23.9%	23.6%	26.5%	26.5%
Mean Burden	1.62	1.44	1.88	1.40	1.25	1.20
Median Burden	1.00	1.00	0.75	0.75	1.00	1.00
Standard Deviation	3.14	1.85	3.95	2.05	1.76	1.29
Standard Error	0.21	0.16	0.31	0.15	0.10	0.16
Importance	**Low**		**Low**		**Low**	
Paid Professional						
Type of Approach (R^2)	***Statistical / Rule Based (.161)***					
Segmentation Variables	*Preparation Method, Primary Form*					
% Impacted	82.4%	82.2%	5.8%	5.8%	5.3%	5.3%
Mean Burden	1.78	1.50	1.44	1.36	3.06	2.97
Median Burden	1.00	1.00	1.00	1.00	2.00	2.00
Standard Deviation	3.63	1.79	3.05	1.73	4.69	4.44
Standard Error	0.12	0.05	0.15	0.12	0.96	0.05
Importance	**Medium**		**Low**		**Low**	
Tax Planning						
Type of Approach (R^2)	***Econometric (.036)***					
% Impacted	38.1%	38.0%	40.8%	40.5%	56.3%	56.2%
Mean Burden	6.85	6.21	6.44	5.27	10.80	8.86
Median Burden	2.00	2.00	2.00	2.00	3.00	3.00
Standard Deviation	16.23	11.31	16.26	9.32	37.26	13.48
Standard Error	0.58	0.52	0.78	0.43	1.30	0.52
Importance	**Medium**		**Medium**		**Medium**	
Form Completion						
Type of Approach (R^2)	***Econometric (.084)***		***Econometric (.227)***		***Econometric (.293)***	
% Impacted	64.0%	63.7%	92.3%	92.2%	97.4%	97.4%
Mean Burden	2.62	2.13	4.02	3.55	5.57	5.22
Median Burden	1.00	1.00	2.00	2.00	3.00	3.00
Standard Deviation	5.63	3.29	7.77	4.98	9.19	6.81
Standard Error	0.22	0.11	0.26	0.15	0.36	0.11
Importance	**Medium**		**High**		**High**	
Form Submission						
Type of Approach (R^2)	***Statistical / Rule Based (.063)***					
Segmentation Variables	*Preparation Method, Form Submission Method*					
% Impacted	32.8%	32.4%	87.5%	87.4%	91.5%	91.5%
Mean Burden	1.05	0.69	1.39	0.98	1.05	0.89
Median Burden	0.33	0.33	0.50	0.50	0.50	0.50
Standard Deviation	3.59	1.11	5.30	1.56	2.29	1.06
Standard Error	0.17	0.04	0.20	0.05	0.11	0.04
Importance	**Low**		**Low**		**Low**	
Total (Outliers Included)						
Population N	45,788,281		35,103,244		10,297,324	
Sample n	2,370		2,329		1,152	

* Weighted data with outliers greater than 5 standard deviations included or removed.

Table B.2
Burden Estimation Framework - Money Burden

	Paid-Prep		Self-Prep w/o SW		Self-Prep w/SW	
	Outliers Included	Outliers Removed	Outliers Included	Outliers Removed	Outliers Included	Outliers Removed
Recordkeeping						
Type of Approach (R^2)	***Statistical / Rule Based (.003)***					
Segmentation Variables	*Primary Form*					
% Impacted	13.7%	13.6%	13.8%	13.7%	18.5%	18.3%
Mean Burden	29.63	27.04	22.16	21.53	29.32	23.10
Median Burden	20.00	20.00	5.00	5.00	10.00	10.00
Standard Deviation	51.71	33.30	49.08	47.64	64.27	34.93
Standard Error	2.65	2.30	4.83	4.83	6.80	2.30
Importance	**Medium**		**Medium**		**Medium**	
Gathering Tax Materials						
Type of Approach (R^2)	***Statistical / Rule Based (.158)***					
Segmentation Variables	*Preparation Method*					
% Impacted	6.9%	6.8%	8.7%	8.7%	57.8%	57.7%
Mean Burden	37.15	36.06	9.63	9.39	38.45	38.15
Median Burden	20.00	20.00	5.00	5.00	30.00	30.00
Standard Deviation	52.26	48.40	12.14	11.38	29.50	28.81
Standard Error	5.45	5.36	1.49	1.49	2.39	5.36
Importance	**Low**		**Low**		**Medium**	
IRS Services						
Type of Approach (R^2)	***Statistical / Rule Based (N/A)***					
Segmentation Variables	*None*					
% Impacted	1.9%	1.9%	0.9%	0.9%	1.0%	1.0%
Mean Burden	30.07	30.07	18.10	18.10	7.52	7.52
Median Burden	19.00	19.00	15.00	15.00	5.00	5.00
Standard Deviation	23.89	23.89	14.11	14.11	6.33	6.33
Standard Error	5.42	5.42	3.31	3.31	2.02	5.42
Importance	**Low**		**Low**		**Low**	
Paid Professional						
Type of Approach (R^2)	***Econometric (.189)***		***Statistical / Rule Based (N/A)***			
Segmentation Variables	*N/A*		*None*			
% Impacted	91.3%	91.3%	4.5%	4.5%	3.0%	3.0%
Mean Burden	113.08	108.47	121.10	114.39	204.22	171.53
Median Burden	85.00	80.00	75.00	75.00	100.00	100.00
Standard Deviation	123.57	92.87	158.29	111.73	321.50	169.46
Standard Error	2.91	2.70	15.77	15.08	52.06	2.70
Importance	**High**		**Low**		**Low**	
Tax Planning						
Type of Approach (R^2)	***Statistical / Rule Based (.002)***					
Segmentation Variables	*Preparation Method*					
% Impacted	3.4%	3.4%	2.3%	2.3%	3.8%	4.3%
Mean Burden	116.76	116.76	38.45	38.45	73.40	73.40
Median Burden	50.00	50.00	20.00	20.00	50.00	50.00
Standard Deviation	209.22	209.22	58.11	58.11	63.41	63.41
Standard Error	38.61	38.61	8.55	8.55	17.02	38.61
Importance	**Low**		**Low**		**Low**	
Form Completion						
Type of Approach (R^2)	***Statistical / Rule Based (.005)***					
Segmentation Variables	*Preparation Method*					
% Impacted	6.0%	6.0%	6.4%	6.4%	8.8%	8.7%
Mean Burden	72.59	70.55	21.32	21.32	40.42	39.57
Median Burden	50.00	50.00	8.00	8.00	20.00	20.00
Standard Deviation	81.25	70.16	35.58	35.58	61.62	60.68
Standard Error	7.25	6.97	3.47	3.47	11.55	6.97
Importance	**Low**		**Low**		**Low**	
Form Submission						
Type of Approach (R^2)	***Statistical / Rule Based (.053)***					
Segmentation Variables	*Preparation Method, Form Submission Method*					
% Impacted	34.2%	34.0%	65.3%	65.1%	74.0%	73.9%
Mean Burden	4.10	3.05	2.62	1.92	6.45	6.28
Median Burden	0.66	0.66	0.57	0.55	2.00	2.00
Standard Deviation	24.07	8.74	8.35	3.98	9.00	8.44
Standard Error	0.68	0.46	0.32	0.15	0.50	0.46
Importance	**Low**		**Low**		**Medium**	
Total (Outliers Included)						
Population N	45,788,281		35,103,244		10,297,324	
Sample n	2,370		2,329		1,152	

* Weighted data with outliers greater than 5 standard deviations included or removed.

APPENDIX C: BURDEN EQUATIONS

Table C.1
Burden Equation Specifications

Dependent Variable: Money Spent on a Paid Professional	Model 1 Paid-Prep	Model 2 Paid-Prep
	5 standard deviation outliers removed	Thompson-Tau outliers removed
Intercept	**60.220**	**58.981**
t statistic	18.141	23.223
FORM VARIABLES		
Schedule A Flag	**30.052**	**33.086**
t statistic	6.714	9.608
Schedule B or 1 Flag	**13.557**	**12.126**
t statistic	2.555	2.983
Schedule D Flag	**48.879**	**36.591**
t statistic	8.956	8.736
Schedule 2 or Form 2441 Flag	**30.087**	**20.762**
t statistic	4.238	3.765
Schedule EITC Flag	**19.939**	**10.570**
t statistic	3.885	2.667
Form 6251 - AMT Flag	**72.045**	**15.447**
t statistic	5.063	1.324
Paid Estimated Taxes	**46.555**	**37.058**
t statistic	5.939	6.029
Form 1116 Flag	**27.845**	**43.213**
t statistic	1.625	3.216
Form 2210 Flag	**43.534**	**18.103**
t statistic	3.411	1.779
Form 8283 Flag	**37.663**	**21.310**
t statistic	2.729	1.975
Other Forms Submitted	**10.908**	**0.170**
t statistic	2.277	0.046
DEMOGRAPHIC VARIABLES		
Interaction among Low AGI, E-Filing, and Refund	**10.070**	**11.339**
t statistic	2.449	3.603
DIAGNOSTICS		
R-Square	**0.189**	**0.184**
F-Statistic Model	**43.249**	**40.649**
Degrees of Freedom	**2231**	**2160**

Table C.2
Burden Equation Specifications

Dependent Variable: Log of Form Completion Time						
	Model 1 Paid-Prep	Model 2 Self-Prep w/o SW	Model 3 Self-Prep w/ SW	Model 4 Paid-Prep	Model 5 Self-Prep w/o SW	Model 6 Self-Prep w/ SW
	5 standard deviation outliers removed			Thompson-Tau outliers removed		
Intercept	**-0.173**	**0.734**	**0.720**	**-0.229**	**0.685**	**0.667**
t statistic	-4.117	21.466	16.281	-5.625	20.925	16.295
FORM VARIABLES						
Owe Taxes	**0.218**	**0.321**	**0.348**	**0.191**	**0.290**	**0.371**
t statistic	2.808	5.727	4.931	2.535	5.344	5.651
DEMOGRAPHIC VARIABLES						
Legal Status: Minor	**-0.374**	**0.098**	**-0.714**	**-0.304**	**0.146**	**-0.666**
t statistic	-2.475	0.627	-5.772	-2.095	0.978	-5.856
Filing Status: Married Filing Jointly	**0.172**	**0.107**	**0.356**	**0.186**	**0.116**	**0.290**
t statistic	2.643	2.072	5.462	2.940	2.344	4.804
Filing Status: Married Filing Separately	**-0.477**	**0.410**	**0.424**	**-0.418**	**0.417**	**0.392**
t statistic	-1.994	2.592	1.932	-1.826	2.728	1.897
TAX SYSTEM ATTRIBUTES						
Attribute Index	**0.031**	**0.091**	**0.090**	**0.032**	**0.085**	**0.095**
t statistic	3.661	11.460	10.330	3.835	11.055	11.605
SUBMISSION METHOD						
Taxpayer submitted return via telefile		**-0.539**			**-0.550**	
t statistic		-8.664			-9.259	
SURVEY VARIABLES						
Spent Time On Forms Not Submitted	**0.770**	**0.451**	**0.428**	**0.792**	**0.404**	**0.440**
t statistic	6.047	7.842	6.276	6.445	7.267	6.892
Multiple Prep Methods (from Survey)	**0.554**	**-0.207**	**0.111**	**0.557**	**-0.151**	**0.124**
t-statistic	6.355	-0.898	1.570	6.609	-0.690	1.874
DIAGNOSTICS						
R-Square	**0.084**	**0.227**	**0.293**	**0.092**	**0.230**	**0.318**
F-Statistic Model	**20.763**	**80.402**	**66.006**	**22.252**	**79.022**	**71.579**
Degrees of Freedom	**1581**	**2191**	**1116**	**1544**	**2115**	**1075**

Table C.3
Burden Equation Specifications

Dependent Variable: Log of Recordkeeping Time	Model 1 Paid-Prep	Model 2 Self-Prep w/o SW	Model 3 Self-Prep w/ SW	Model 4 Paid-Prep	Model 5 Self-Prep w/o SW	Model 6 Self-Prep w/ SW
	5 standard deviation outliers removed			Thompson-Tau outliers removed		
Intercept	**0.755**	**0.667**	**0.673**	**0.689**	**0.583**	**0.600**
t statistic	16.731	14.873	10.325	15.831	13.758	9.858
CHANGE VARIABLES (TY 98 to TY 99)						
New Zip Code (TY 98 to TY 99)	**-0.058**	**0.268**	**0.264**	**-0.029**	**0.282**	**0.055**
t statistic	-0.633	3.532	2.521	-0.325	3.940	0.546
DEMOGRAPHIC VARIABLES						
Legal Status: Minor	**-0.505**	**-0.349**	**-0.343**	**-0.431**	**-0.335**	**-0.265**
t statistic	-3.229	-1.692	-2.016	-2.885	-1.729	-1.687
Filing Status: Married Filing Jointly	**0.207**	**0.230**	**0.645**	**0.181**	**0.180**	**0.603**
t statistic	3.051	3.290	6.795	2.761	2.714	6.715
Filing Status: Married Filing Separately	**-0.627**	**0.534**	**0.505**	**-0.599**	**0.513**	**0.616**
t statistic	-2.757	2.608	1.494	-2.747	2.643	1.979
TAX SYSTEM ATTRIBUTES						
Attribute Index	**0.071**	**0.088**	**0.058**	**0.067**	**0.086**	**0.058**
t statistic	8.000	8.724	4.734	7.833	9.023	4.932
SURVEY VARIABLES						
Spent Time On Forms Not Submitted	**0.388**	**0.407**	**0.281**	**0.428**	**0.386**	**0.244**
t statistic	2.787	4.984	2.896	3.188	4.952	2.655
Multiple Prep Methods (from Survey)	**0.585**	**0.442**	**0.266**	**0.547**	**0.523**	**0.326**
t statistic	5.998	1.372	2.654	5.780	1.737	3.468
DIAGNOSTICS						
R-Square	**0.082**	**0.093**	**0.154**	**0.080**	**0.099**	**0.160**
F-Statistic Model	**24.347**	**29.607**	**27.977**	**22.892**	**30.000**	**27.917**
Degrees of Freedom	**1914**	**2014**	**1076**	**1855**	**1923**	**1026**

Table C.4
Burden Equation Specifications

Dependent Variable: Log of IRS Services Time	Model 1 Paid-Prep	Model 2 Self-Prep w/o SW	Model 3 Self-Prep w/ SW	Model 4 Paid-Prep	Model 5 Self-Prep w/o SW	Model 6 Self-Prep w/ SW
	5 standard deviation outliers removed			Thompson-Tau outliers removed		
Intercept	**0.247**	**0.193**	**-0.084**	**0.288**	**0.140**	**-0.094**
t statistic	1.201	1.763	-0.484	1.459	1.443	-0.555
FORM VARIABLES						
Form 1040A Flag	**-0.036**	**-0.006**	**-0.595**	**-0.006**	**-0.075**	**-0.497**
t statistic	-0.275	-0.073	-4.023	-0.048	-0.988	-3.494
Form 1040EZ Flag	**0.164**	**-0.033**	**-0.599**	**0.178**	**-0.013**	**-0.724**
t statistic	0.776	-0.336	-4.832	0.888	-0.151	-5.861
SURVEY VARIABLES						
Visited Walk-In Sites for Tax Assistance	**0.492**	**0.354**	**0.412**	**0.457**	**0.411**	**0.335**
t statistic	2.828	3.664	2.163	2.715	4.742	1.781
Called Tele-Tax to Hear Message About Tax Law	**0.117**	**0.259**	**0.493**	**0.104**	**0.307**	**0.350**
t statistic	0.845	3.335	4.018	0.782	4.460	2.810
Log of IRS Services Wait Time	**0.683**	**0.603**	**0.430**	**0.706**	**0.597**	**0.414**
t statistic	13.151	19.379	9.165	13.475	21.520	8.929
Respondent called the IRS Toll Free Tax Assistance line	**0.142**	**0.235**	**0.296**	**0.103**	**0.188**	**0.240**
t statistic	0.958	2.878	2.453	0.720	2.609	2.053
Respondent used IRS Web site for other than form download	**0.738**	**0.367**	**0.707**	**0.731**	**0.419**	**0.742**
t statistic	4.076	3.987	6.002	4.234	5.174	6.493
DIAGNOSTICS						
R-Square	**0.595**	**0.550**	**0.531**	**0.610**	**0.605**	**0.542**
F-Statistic Model	**34.468**	**70.419**	**28.136**	**35.450**	**87.189**	**27.888**
Degrees of Freedom	**164**	**404**	**174**	**159**	**398**	**165**

Table C.5
Burden Equation Specifications

Dependent Variable: Log of Gathering Tax Materials Time	Model 1 Paid-Prep	Model 2 Self-Prep w/o SW	Model 3 Self-Prep w/ SW	Model 4 Paid-Prep	Model 5 Self-Prep w/o SW	Model 6 Self-Prep w/ SW
	5 standard deviation outliers removed			Thompson-Tau outliers removed		
Intercept	**-0.232**	**-0.490**	**0.043**	**-0.284**	**-0.528**	**-0.009**
t statistic	-5.092	-12.135	1.149	-6.426	-13.405	-0.263
FORM VARIABLES						
Paid Estimated Taxes	**0.350**	**0.375**	**0.031**	**0.338**	**0.350**	**-0.011**
t statistic	2.412	2.279	0.181	2.394	2.178	-0.069
SUBMISSION METHOD						
Taxpayer submitted return via telefile		**-0.657**			**-0.700**	
t statistic		-8.138			-8.920	
SURVEY VARIABLES						
No. of Times Visited Library, Post Office, etc.	**0.134**	**0.305**	**0.073**	**0.139**	**0.321**	**0.100**
t statistic	2.881	10.543	1.359	3.097	11.354	1.938
No. of Times Ordered/Downloaded Forms/Pubs from IRS	**0.064**	**0.220**	**0.097**	**0.076**	**0.222**	**0.106**
t statistic	0.515	5.500	2.160	0.635	5.545	2.466
Obtained Books or Guides For Tax Purposes	**0.575**	**0.534**	**0.470**	**0.516**	**0.525**	**0.343**
t statistic	2.692	4.382	3.324	2.468	4.359	2.449
Multiple Prep Methods (from Survey)	**0.238**			**0.290**		
t-statistic	2.172			2.749		
DIAGNOSTICS						
R-Square	**0.026**	**0.155**	**0.020**	**0.030**	**0.171**	**0.019**
F-Statistic Model	**5.552**	**58.182**	**4.783**	**6.238**	**63.967**	**4.366**
Degrees of Freedom	**1039**	**1585**	**918**	**1023**	**1556**	**895**

Table C.6
Burden Equation Specifications

Dependent Variable: Log of Tax Planning Time	Model 1 All Prep Methods	Model 2 All Prep Methods
	5 standard deviation outliers removed	Thompson-Tau outliers removed
Intercept	**0.739**	**0.619**
t statistic	15.118	13.292
PREPARATION METHOD		
Self-Prep w/o SW	**-0.097**	**-0.058**
t statistic	-1.834	-1.145
Self-Prep w/ SW	**0.327**	**0.411**
t statistic	4.789	6.379
DEMOGRAPHIC VARIABLES		
Filing Status: Married Filing Jointly	**-0.134**	**-0.067**
t statistic	-2.364	-1.255
Filing Status: Married Filing Separately	**-0.374**	**-0.265**
t statistic	-2.377	-1.797
Disposable Income in 1000s	**0.007**	**0.006**
t statistic	5.382	5.221
Disposable Income in 1000s Squared	**-0.000**	**-0.000**
t statistic	-2.638	-2.485
TAX SYSTEM ATTRIBUTES		
Attribute Index	**0.009**	**0.010**
t statistic	1.124	1.324
DIAGNOSTICS		
R-Square	**0.036**	**0.045**
F-Statistic Model	**17.121**	**20.879**
Degrees of Freedom	**3182**	**3114**